

MEMORANDUM
RM-4196-ARPA
JANUARY 1966

AD630818

A CRITICAL REVIEW OF
THE EXPERIMENTAL FOUNDATION OF
HUMAN COLOR PERCEPTION

Joseph J. Sheppard, Jr.

CLEARINGHOUSE FOR FEDERAL SCIENTIFIC AND TECHNICAL INFORMATION			
Hardcopy	Microfiche		
\$ 5.00	\$ 1.00	187 pp	00
ARCHIVE COPY			

Code 1

PREPARED FOR:

ADVANCED RESEARCH PROJECTS AGENCY

The RAND Corporation
SANTA MONICA • CALIFORNIA

MEMORANDUM
RM-4196-ARPA
JANUARY 1966

A CRITICAL REVIEW OF
THE EXPERIMENTAL FOUNDATION OF
HUMAN COLOR PERCEPTION

Joseph J. Sheppard, Jr.

This research is supported by the Advanced Research Projects Agency under Contract No. SD-79. Any views or conclusions contained in this Memorandum should not be interpreted as representing the official opinion or policy of ARPA.

DISTRIBUTION STATEMENT

Distribution of this document is unlimited.

The RAND Corporation
1700 MAIN ST. • SANTA MONICA • CALIFORNIA • 90406

Approved for release by the Clearinghouse for
Federal Scientific and Technical Information

BLANK PAGE

PREFACE

Physical scientists and engineers have become increasingly concerned with the potential utilization of inherent characteristics of biological organisms as a means of solving complex technological problems. One such characteristic of widespread interest is the human ability to discriminate subtle differences in perceived colors.

However, in any proposed employment of human color perception a new problem presents itself: How is one to distinguish between hypothesis and established experimental fact? This problem is acute for the relative newcomer to the subject. On the one hand he is faced with a few cyclopedic treatments of particular details of the visual process containing much more specific information than he requires in his feasibility studies. As an alternative he finds a plethora of textbook-style reviews typically containing brief summaries of a few experimental findings collected in support of, or presented within the framework of, a particular theory of the visual mechanism. Neither type of treatise satisfies the need for a critical presentation of the minimum material necessary for a comprehensive view of the subject. The present work is an attempt at partial fulfillment of this need.

Hence, this study does not purport to be either a general compendium or a detailed discussion of some specific aspect of color vision. Rather, it is intended to be a syllabus, presented through an ordered discussion of some important available experimental results. Individual data and groups of related data are reanalyzed for their inherent significance without regard to whether they support, refute, or remain neutral toward any specific "theory" of the visual mechanism. The results of these analyses are then formulated into experimentally based conclusions. Finally, the principal conclusions are discussed with a view toward their contributions to the general understanding of normal human color perception.

The present study also differs sharply from most discussions of human color perception in that no specific attention is given either to extant theories of the visual mechanism or to color vision anomalies. Available theories are excluded, since the present purpose is an

interpretation of experimental data for their intrinsic content. The considerable data on subjects with reduced color vision have been excluded as beyond the scope of a treatise on normally functioning systems. A comprehensive familiarity with the manifold characteristics of normal human color perception is a mandatory precursor to fruitful consideration of color vision anomalies.

The study reported in this Memorandum is a part of an investigation for ARPA of the possible utility of a man in a missile discrimination system.

SUMMARY

A working definition is presented of the overall experimental problem in the study of normal human color perception. The artificial nature of colorimetry is discussed, with particular emphasis on the distinction between the experimental facts established in the matching experiments and the formalism of colorimetry derived in part from these facts.

A representative portion of the available experimental data on individual foveal spectral sensitivity is collected and analyzed. A major conclusion is that any attempt to fit photopic luminosity data with three or four fundamental cone mechanisms, each with a fixed spectral response curve, shall very probably fail of general usefulness.

Experimental data are analyzed concerning the initial photoreception process and the anatomy, histology, morphology, ontogeny, and electrophysiology of the retinal neurons. Attention is drawn to the multiplicity of results indicating a fundamental difference between receptor mechanisms in the rods and cones. The major conclusions of these analyses imply that the peripheral mechanism subserving human color vision is not based simply on three cone photopigments, but rather is based on a combination of variation in cone physical characteristics (including any pigments) and multiple bipolar analyses. This implies that the Young-Helmholtz three-pigment theory, the Hering opponent-colors theory, and the various stage or zone theories may each contain some element of truth, but that none is individually capable of providing a suitable explanation.

A review of data on central neural mechanisms indicates a complex, dynamic role for the lateral geniculate nuclei in human color vision. Further, it is concluded that the neural signal complex synapsing in the cerebral cortex is not afferent determinate, but is under the dynamic control of still higher brain centers.

Considerations of rather diverse psychophysiological phenomena are summarized into three major conclusions: (1) In its dynamic aspects, the human visual system displays a continuous effort to reset its various thresholds to a pattern whereby it is most sensitive to deviations

from a current "equilibrium" level. (2) Insofar as color vision is concerned, the stimulus considered as white (or achromatic) is the principal determinant for this equilibrium level. The actual spectral distribution of the "white" stimulus is important, not just its CIE chromaticity. (3) In a dynamic utilization of the preceding two features, the central nervous system always attempts to perceive order in the afferent signals consistent with past experience. On finding even a minimal content satisfying such an order, it will attempt to reject the disordered content and accept and enhance the ordered content, to enable the perception of familiar sensations.

The general conclusion reached in this study is that the available experimental evidence does not clearly dictate the fundamental physiological processes mediating human color vision.

The principal conclusions are discussed relative to their implications for the three distinct fields of colorimetry, visual biophysics, and visual psychophysics. Four postulates are recommended for more fruitful psychophysical modeling: (1) The specification of stimuli only in terms of CIE coordinates should be supplemented (or replaced) by specifications which describe the spectral distribution of a stimulus. (2) The ability of a stimulus to produce a color sensation should be measured in terms of its deviation from some equilibrium stimulus, and the actual sensation should be causally related to the details of the deviation. (3) The equilibrium stimulus should be related to the stimulus currently perceived as white or achromatic, but allowance must be made for simultaneous "effective" equilibrium stimuli, all of which need not be consciously perceived by the subject. (4) A successful model must be able to predict more than one sensation for a given input (i.e., a different sensation for each of the currently active "effective equilibrium stimuli"), each prediction thus representing a possible perception, depending on the observer's intention.

ACKNOWLEDGMENTS

The author wishes to express his sincere appreciation to Carl Gazley of The RAND Corporation for his sustained interest, helpful discussions, and many useful suggestions during the preparation of this study.

Constructive criticisms of the manuscript provided by Doris Dugas, Carl Gazley, Sam Genensky, Hubert Moshin, Roy Stratton, and Donna Wilson are gratefully acknowledged.

The author is also very grateful to Kathleen Blackwood, Jeannine McGann, and Hubert Moshin for assistance with the many illustrations.

BLANK PAGE

CONTENTS

PREFACE	iii
SUMMARY	v
ACKNOWLEDGMENTS	vii
SYMBOLS	xi
Section	
I. INTRODUCTION	1
II. THE PROCESS OF COLOR PERCEPTION	4
The Visual System	4
The Domains of Color Vision	7
III. THE STANDARD OBSERVER	11
The CIE Brightness Response	11
Metamers	14
The Matching Experiments	17
Trivariance	21
Chromaticity	23
The CIE Chromatic Response	27
The Artificiality of the CIE Response	36
IV. RETINAL SENSITIVITY	38
Retinal Duality	38
Relative Sensitivity	39
Scotopic Sensitivity	41
Photopic Sensitivity	51
Conclusions	64
V. RETINAL PHYSIOLOGY	66
Retinal Photoreception	66
Retinal Anatomy	76
Receptor Histology	83
Receptor Differentiation	90
Retinal Electrophysiology	93
Conclusions	101
VI. CEREBRAL PHYSIOLOGY	103
Subcortical Synapses	105
Visual Cortex	114
VII. PSYCHOPHYSIOLOGY	117
Adaptation	117
The Yellow Anomaly	127
The White Anomaly	132
The Kohler and Land Experiments	133
Summary	149

VIII. CONCLUSIONS AND RECOMMENDATIONS	152
Summary of Principal Conclusions	152
Discussion and Recommendations	156
APPENDIX	159
Radiant Energy Terminology	159
Luminous Energy Terminology	160
Photometric Units	162
REFERENCES	165

SYMBOLS

- F = luminous flux
- F_{λ} = luminous flux per unit wavelength interval about λ
- K = luminous efficiency (678.8 lm/watt)
- L = luminance
- L_a = luminance level corresponding to state of brightness adaptation
- L_b = luminance level at "black point"
- $L_i(\lambda_{ci})$ = luminance of a "pure purple" light, defined by Eqs. (8) and (9)
- $L_i(\lambda_i)$ = L_i luminous units of light at wavelength λ_i
- L_w = luminance of an equal energy continuous spectrum "white"
- n = index of refraction
- P = radiant flux
- P_{λ} = radiant flux per unit wavelength interval about λ
- P_i = arbitrarily defined unit of luminance for a given primary
- p_e = excitation purity, defined by Eq. (22)
- V_{λ} = CIE 1924 Photopic Luminous Efficiency Function
- V'_{λ} = CIE 1951 Scotopic Luminous Efficiency Function
- X, Y, Z = CIE Standard (imaginary) Primaries for colorimetry
- (x, y, z) = chromaticity coordinates of a light stimulus in the XYZ System
- $(\bar{x}, \bar{y}, \bar{z})$ = distribution coefficients, i.e., the tristimulus values of the spectrum in the XYZ System for the CIE Standard Observer
- $\Gamma_i(P_i)$ = tristimulus value, i.e., the amount of the unit quantity P_i

Y_i = trichromatic coordinate, defined by Eq. (16)

Λ = any visible electromagnetic stimulus, monochromatic or continuous

λ = wavelength; dominant wavelength

λ_c = complimentary wavelength

Note: The symbol \equiv is to be read as "matches" or "is equivalent in sensation to."

I. INTRODUCTION

There are enormous and ever-increasing amounts of experimental data concerning the physical, biological, and psychological aspects of color perception. (See, for example, Refs. 1-18.) Any proposed theory of human color perception must take cognizance of this information. At the same time, the limitations of such data must be recognized in the sense of their incompleteness and frequent lack of self-agreement. Seemingly their tendency is toward a concentration on minute details, with only secondary concern for any generalized concepts of the overall sensory process. From the writer's purview, the fundamental problem extant in human color vision is precisely the lack of a generalized theory that is (1) cognizant of, but not irrevocably dependent on, the presently accepted details; (2) developed from a minimum of reasonable postulates; and (3) capable of producing experimentally verifiable predictions over a broad range of practical situations without additional arbitrary assumptions peculiar to each situation.

Such an attitude in no way implies a slighting of the value of experimental work on any detailed aspect of the visual process. To the contrary, the intent is to emphasize (in the absence of such a generally acceptable theory) the a priori equality of importance of all such data. By equality is meant that all data, both direct and indirect, should be considered for their possible contributions to the general solution. Unfortunately, the preponderance of the data on human color vision is open to several interpretations. It is probably this crucial fact that has to date precluded a satisfactory general solution and, at the same time, has continued to entice new theorists into the field.

Now the only immediately apparent feature characterizing the data on human color vision is the rather large variation from subject to subject in response to a supposedly constant set of stimuli. Indeed, this seems to be a rule for biological processes. In synthesizing experimental results obtained by varying procedures and on different subjects, one should therefore look for correlates in the form of deviations from individual base lines and should refrain from attaching too much significance to individual absolute values or averages thereof.

Such an approach may imply a prominent role for statistics in any theoretical treatment. This would not necessarily be artificial, since the visual process is in a sense a communications system between the organism and its surround; and the central problem⁽¹⁹⁾ of all communications systems is essentially one of statistical inference. (Shall one equate perception with statistical inference?)

Furthermore, in any theoretical development purporting to correlate the experimental results, one should avoid the tendency to concentrate unduly on peripheral mechanisms, with little or no regard to cortical processes. The very concept of color sensation is devoid of meaning except in a perceptual (i.e., cortical) sense. The writer feels that only by an initial inclusion of appropriate cortical hypotheses with the peripheral ones, can a theory be developed that will describe with any reasonable degree of success the changeable aspects of human color vision. Finally, for a theory so constructed to be worthy of serious consideration, it must be shown that the postulates advanced are both plausible and compatible with the experimentally established physical, psychological, and physiological facts of the visual process.

But what are these facts? The study of human color vision involves a very large number of apparently disparate problems, from the physics and chemistry of the photoreceptive process to the subjective insistence of yellow as a fundamental hue. The interrelationship of the three scientific disciplines principally involved is illustrated in Fig. 1. Each subdivision of this illustration represents an encyclopedia of information pertinent to various aspects of color vision, and it would be a forbidding task to undertake a detailed examination of even one of them. At the other extreme, the validity of many an extant "theory" of color vision is based on an appeal to a very few experimental facts principally drawn from a single discipline. Both the assessment of applicability of such theories and proper development of alternative hypotheses dictate the pursuit of some middle course.

Many summaries of experimental findings pertinent to color vision are available. Some of these are listed as Refs. 1-18. They range, for example, from the general reviews by LeGrand⁽²⁾ and Osgood⁽¹⁷⁾ to the details of anatomy by Polyak^{(8),(9)} and of peripheral mechanisms

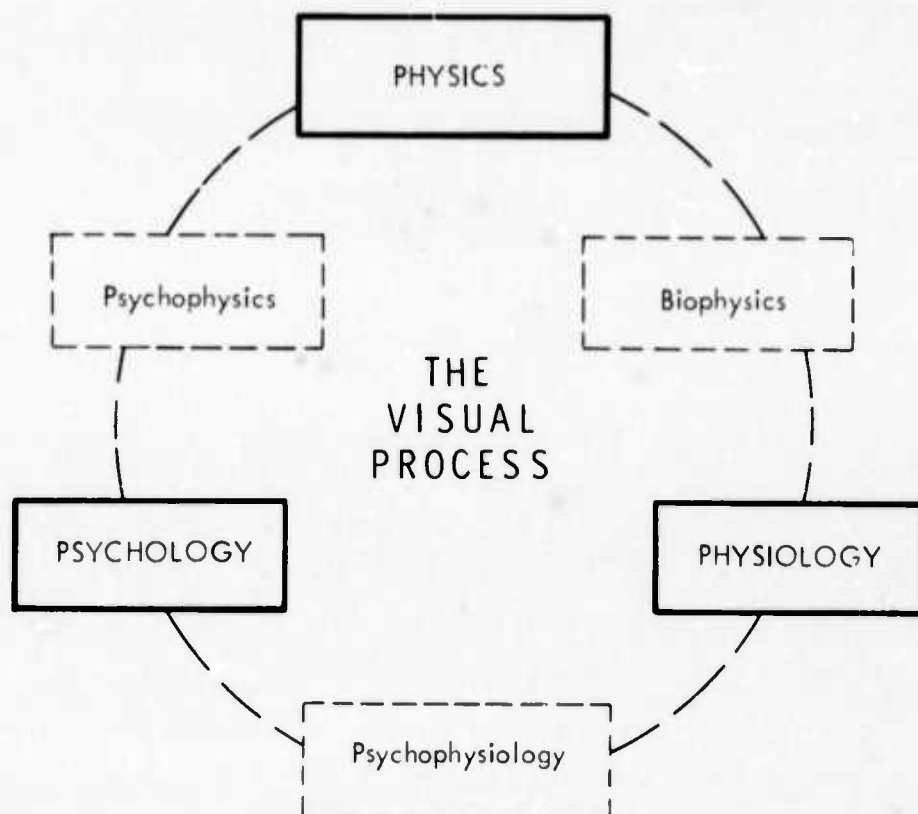


Fig.1 — Scientific disciplines involved in the study of the visual process

by Talbot. (18) The present study does not purport to be either a general compendium or a detailed discussion of some specific aspect of color vision. Rather, it is intended to be a syllabus, presented through an ordered discussion of some important available experimental results. These results are discussed here with a view toward their contributions to the general solution of the problems of human color perception.

In view of the diversity of topics discussed and the varying backgrounds of the probable audience, a certain amount of introductory material is included in each section, for continuity and intelligibility. The reader who is familiar with colorimetry as embodied in the CIE* Standard Observer is forewarned that Section III is not a standard presentation of the CIE System, but, rather, stresses the essential artificiality of this widely misunderstood formalism.

*Commission Internationale de l'Eclairage (International Commission on Illumination).

II. THE PROCESS OF COLOR PERCEPTION

THE VISUAL SYSTEM

The various stages in the overall process of color perception in man may be described in deceptively simple terms. As shown in Figs. 2 and 3, electromagnetic radiation reflected from the "red" bird is focused by the eye on the retinal receptors (rods and cones), whence, under appropriate conditions, a system of nerve currents is transmitted to the higher centers of the brain, resulting in the perception of form (bird) and color (redness). Stage 1 of this process is concerned with specification of the spectral composition of the radiation reflected from the bird and its subsequent modification by various parts of the eye (cornea, lens, etc.) before final incidence on the photosensitive parts of the receptors. In Stage 2 some of the radiation is absorbed in the receptor neurons. This absorption initiates a neuroretinal process that culminates in the transmission of electrical impulses (afferent signal complex) along the optic nerve fibers toward the brain proper. In Stage 3, at the chiasma, the optic nerve bundles from each eye meet; thence some fibers from the right eye cross over (decussate) to the left optic tract, and some fibers from the left eye cross over to the right optic tract. There are no electrical connections (synapses) at the chiasma. There is, however, a rearrangement of the nerve fibers from the two incoming bundles, one from each eye, into two outgoing bundles each containing fibers from both eyes. In Stage 4, the retinal nerve fibers terminate (synapse) in the subcortical visual centers, principally in the lateral geniculate body. Finally, cortical nerve fibers originate in the lateral geniculate nuclei (LGN) and terminate in the striate areas of the occipital lobes (visual cortex). This constitutes Stage 5 of the visual process.

The deceptive nature of the preceding description becomes apparent when it is realized that the mechanism for the initial radiation reaction in the cone receptors is still unknown, to say nothing of the nature of the final emergence of a color sensation in human consciousness. For example, radiation of wavelength 650 millimicrons ($m\mu$) may be called "red light" by the physicist, but this is not a specification

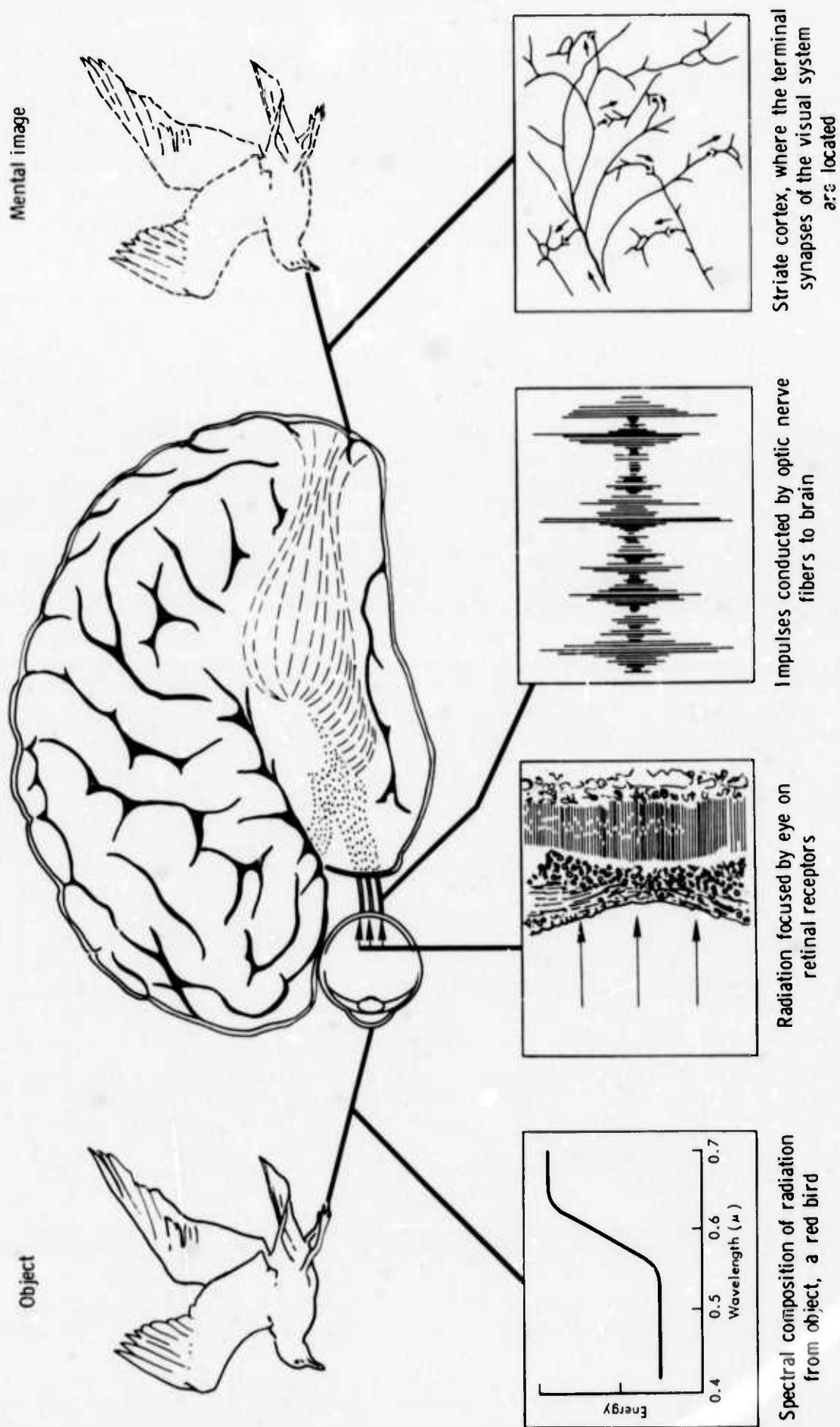
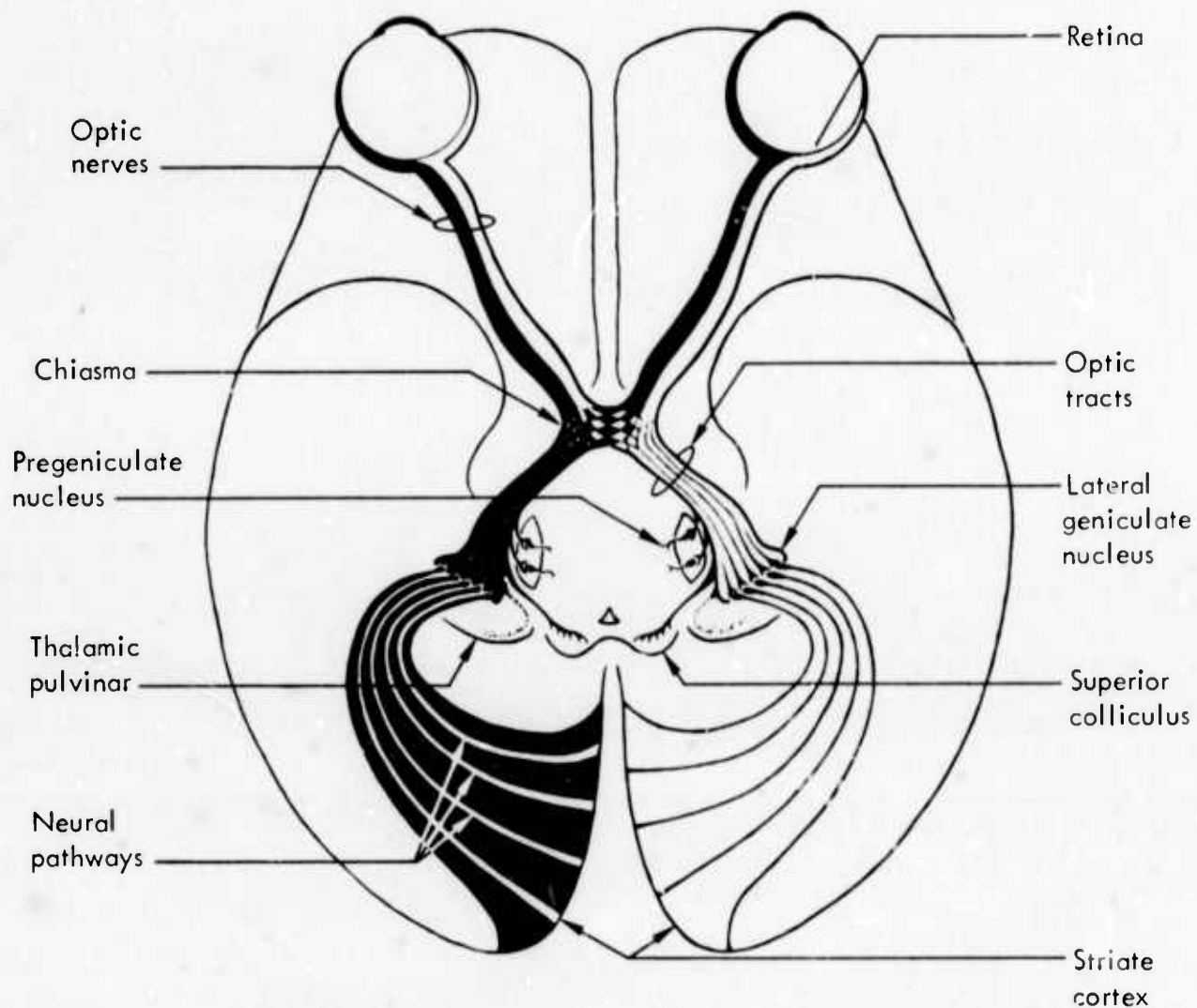


Fig.2 —Stages of the visual process (adapted from Ref. 20)



- The subcortical portion consists of the retinas, optic nerves, chiasma, and optic tracts, the latter terminating in subcortical visual centers, principally the lateral geniculate nucleus, with subsidiary pregeniculate nucleus, superior colliculus, and thalamic pulvinar
- The supranuclear division consists of the neural pathways originating in the lateral geniculate nuclei and terminating in the striate areas of the occipital lobes
- Fibers originating in the inner or nasal halves of the retinas inter-cross in the chiasma

Fig.3—Diagram of the human visual system drawn into an outline of the brain (adapted from Ref. 9)

of the psychological sensation of redness. Indeed, while the radiation of 650 mμ may be perceived as a "red" most of the time by most people, it is also true that under appropriate conditions other radiation containing no energy at 650 mμ is indistinguishable perceptually from the 650-red. Under still different conditions the 650-radiation may be perceived as some color other than red. In short, the human visual system does not behave as an inorganic instrument, but rather is, in operation, a dynamic, rapidly changing, extremely complex biological system.

Such a situation constitutes a continuing challenge to the many investigators of color vision, but it is an imprecation to those involved with color technology. The latter must have some reference model, a universal "standard human observer," whose fundamental requisite is that its "perceptual" reaction to any given radiant stimulus is completely defined. A desirable but secondary feature of a standard observer is that its "response" to a given stimulus should be "typical" of those of actual individuals. Such a model was provisionally defined in 1924 by the CIE. The international standards of the CIE have been embodied in a monumental project of the Committee on Colorimetry of the Optical Society of America in the creation of a major subdivision of the study of human color vision, the psychophysics⁽²¹⁾ of color. The characteristics of this model (hereinafter called the "Standard Observer") and the underlying experimental data are the subject of Section III. But first a few words are in order concerning the subdisciplines mentioned in the Introduction.

THE DOMAINS OF COLOR VISION

With reference to human vision, the term "color" may be considered to have four distinct meanings, the one implied being determined by which domain⁽³⁾ of the overall visual process is under discussion. Three of these domains--the physical, the physiological, and the psychological--are "real" in the sense that each represents an assemblage of experimental facts. In the first two, purely objective procedures are used, and the latter uses primarily subjective techniques. The interrelationship of these three domains is indicated in Fig. 4.

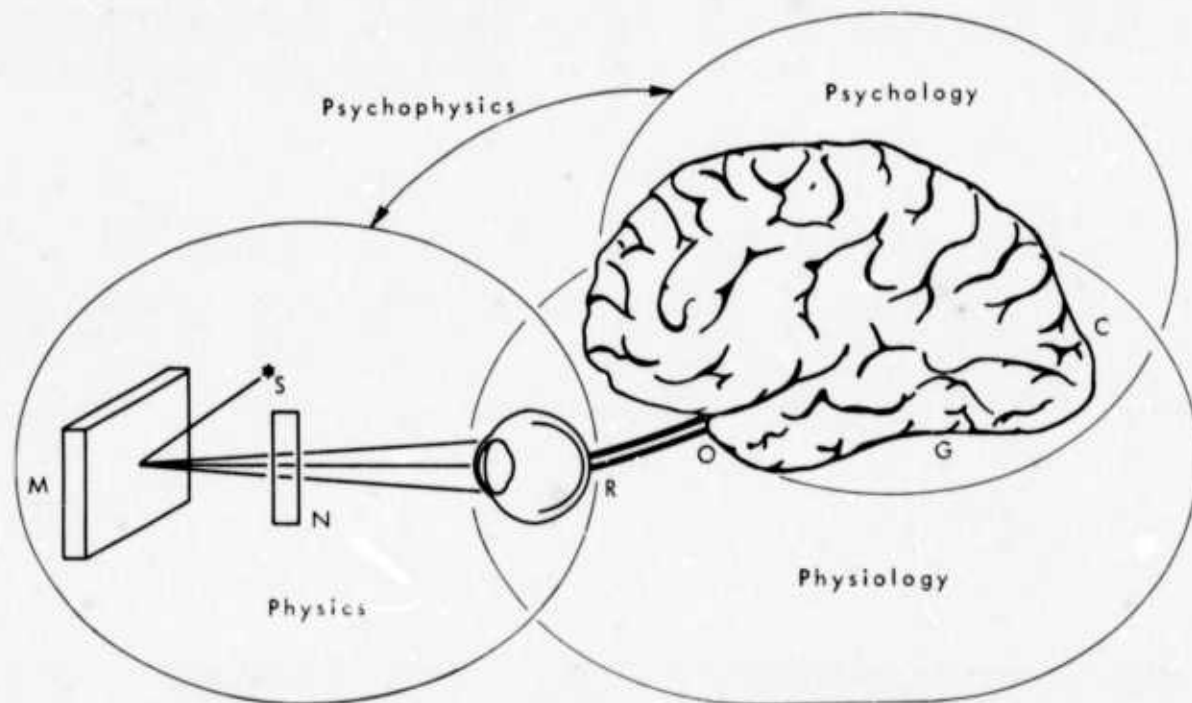


Fig.4— Interrelationship of the three domains of the visual process
(adapted from Ref. 1)

The physicist can specify with reasonable accuracy the spectral intensity of the radiation from source S incident on the retina R after the radiation passes through the monochromator M and neutral filter N. Thus, from the physicist's viewpoint the term "spectral composition" of the radiation is to be preferred to the term "color."

The physiologist is concerned with the as yet unsolved problem of determining the "color frequency-code" of the nerve impulses propagating along the optic nerve O to the subcortical areas G and thence to the cerebral cortex C.

Psychological evidence indicates that the arrival of an afferent signal at G does not necessarily result in the conscious perception of a sensation at C unless the psyche so desires (i.e., is attentive). If a sensation is perceived upon incidence of the radiation on the retina, then the subject may so indicate by describing the sensation or by reporting when two alternate sensations are indistinguishable. Thus the preferred psychological term corresponding to the radiation is "color sensation."

To satisfy a practical need in engineering applications, one may postulate a causal relationship between the incident physical stimulus and the psychological response, ignoring the physiological link. Psychophysics,⁽²²⁾ the fourth visual domain, is concerned with such relationships as can be discovered between physical stimuli and the perceptual aspects of the responses to these stimuli. However, it should be noted that the psychophysical domain is essentially artificial. Its subject matter is obtained by defined generalizations, in physical terms, of concepts derived from the subjective aspects of the response of the human organism to physical stimuli.

Some of the terminology in the psychophysics of vision is shown in Fig. 5, along with corresponding terms in physics and psychology.* This standardization results in the unqualified word "color" having a well-defined meaning only in the psychophysical or theoretic plane. There it is normally determined by the formalism of colorimetry^{(21), (23)} in terms of a defined "standard observer." It is important to note that whereas the particular color name assigned by colorimetry to radiation of a given spectral composition would be acceptable to many people and might be entirely adequate for most industrial purposes, the fact remains that the color so determined has no intrinsic connection with the color sensation perceived when the given radiation is incident on the retina of a particular human observer. This artificial aspect of the psychophysical terms must be kept clearly in mind in any discussion of the physical facts of the visual process, since the latter are frequently presented in psychophysical language. Great care is often required to distinguish between proper experimental fact and theoretic generalization.

* A more recent effort in terminology standardization is presented in Ref. 6.

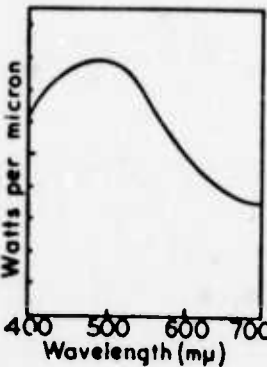
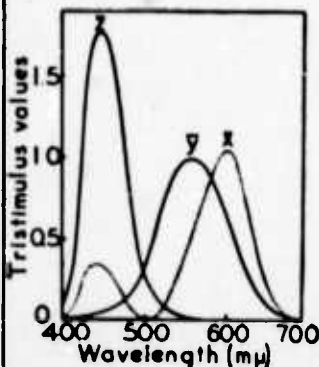
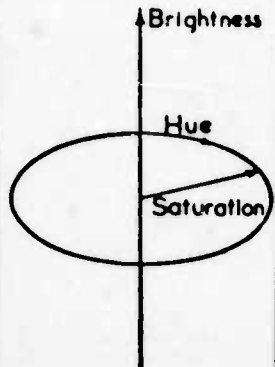
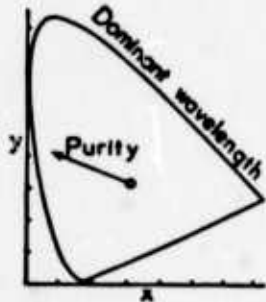

PHYSICS	PSYCHOPHYSICS	PSYCHOLOGY	
Visual Stimulus →	Light	← Visual Sensation	Visual Perception
Radiant energy Spectral composition →	Luminous energy Color	← Color sensation	
			
Characteristics of Radiant Energy	Characteristics of Light = Color	Attributes of Color Sensation	Corresponding Modes of Appearance
Radiant flux →	Luminous flux	← Brightness	Aperture (1-5)
Radiance →	Luminance		Illuminant (1-8)
Irradiance →	Illuminance		Illumination (1-3)
Radiant reflectance →	Luminous reflectance		Object modes: Surface (1-11)
Radiant transmittance →	Luminous transmittance		Volume (1-9)
Spectral distribution →	Chromaticity	← Chromaticness	Attributes of modes of appearance:
(Relative spectral composition, quality)	Dominant wavelength (or complementary)	Hue	1. Brightness (or lightness)
Radiant purity	Purity	Saturation	2. Hue
			3. Saturation
			4. Size
			5. Shape
			6. Location
			7. Flicker
			8. Sparkle
			9. Transparency
			10. Glossiness
			11. Luster

Fig.5—Nomenclature in color vision (adapted from Ref. 22)

III. THE STANDARD OBSERVER

THE CIE BRIGHTNESS RESPONSE

The human eye is sensitive to only a narrow band of electromagnetic radiation, known as the visible spectrum, whose limits depend on the intensity of the radiation involved. These limits define what is meant by the psychophysical term "light." Radiation near the middle of the visible spectrum is most effective in eliciting the visual brightness response, the effectiveness decreasing toward the long and short wavelength limits. The curve representing this sensitivity is called a "luminous efficiency function" or "luminosity function" and may be obtained for any given observer in a variety of ways. Because it expresses the relative abilities of physical stimuli of different wavelengths to elicit the psychological sensation of brightness, the luminosity function is obviously a psychophysical concept.

A determination of such sensitivity curves at widely different illumination levels has established the fundamental duality of the human retina. There are two distinct types of retinal receptors: rods, which function in dim light (scotopia) where the sensations perceived are achromatic, i.e., colorless; and cones, the receptors active in bright light (photopia) where color vision is fully developed. The corresponding CIE-defined luminosity functions are shown in Fig. 6 as the log of the reciprocal energy (relative to the maximum) incident at the cornea. The maximum of each curve has been arbitrarily set at unity so that there is no indication of the comparative sensitivities of the two receptor systems. The enormous dynamic range of the human visual system between the rod threshold and the glare limit is presented schematically in Fig. 7, which shows a ratio of perceptible intensities of 10^{10} . The transition from scotopic to photopic vision is not sudden but gradual over the approximate range of 0.001 to 0.1 millilambert (mL). Since Fig. 6 shows that the scotopic curve peaks at about 507 mμ, while the photopic function has a maximum at about 555 mμ, in the transition region (mesopia) from dim to bright light the spectral sensitivity maximum of the eye shifts toward the red. This is called the "Purkinje

Note:

The official tabulations of these functions by the CIE⁽²⁴⁾ are virtually inaccessible. The most convenient source for both functions is Ref. 2, page 111 for V'_λ and pages 454-455 for V_λ , the latter being tabulated as \bar{y} .

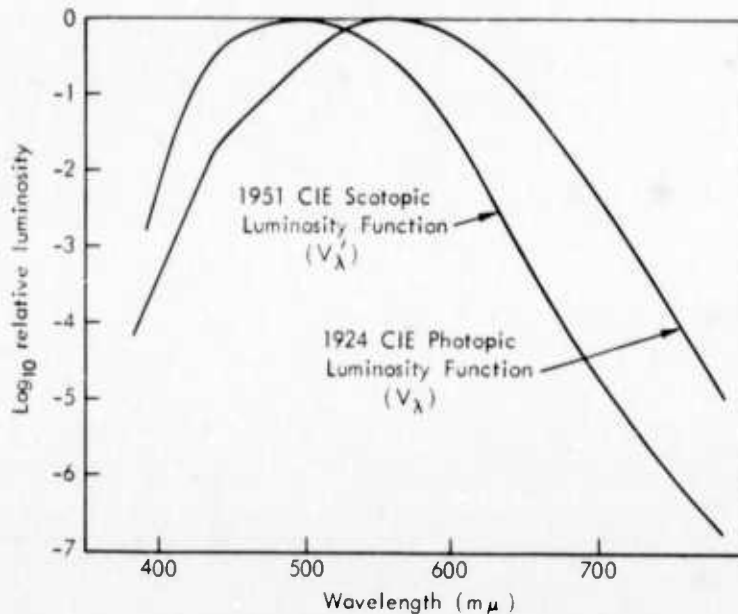


Fig.6—CIE Standard Luminous Efficiency Functions for scotopic and photopic vision (Ref. 24)

Effect" and may sometimes be noticed in late twilight when green leaves assume an unnatural brightness relative to yellow and red objects. Since the scotopic function peaks in the green, whereas the photopic maximum is in the yellow region of the spectrum, if green and yellow objects are seen as having the same brightness in daylight, the green will appear brighter in twilight (although both will be gray).⁽³⁾ The CIE makes no provision for the Purkinje Effect, and hence there are no standard procedures available for photometric measurements in the mesopic region.

With a defined standard observer replacing the variable human, an entire system of psychophysical or "luminous" quantities may be developed from the purely radiant ones. If the CIE 1924 Photopic Luminous Efficiency Function is denoted by V_λ , then the fundamental relation between the psychophysical quantity called "luminous flux," F , and the purely physical "radiant flux," P , is

$$F = K \int V_\lambda P_\lambda d\lambda. \quad (1)$$

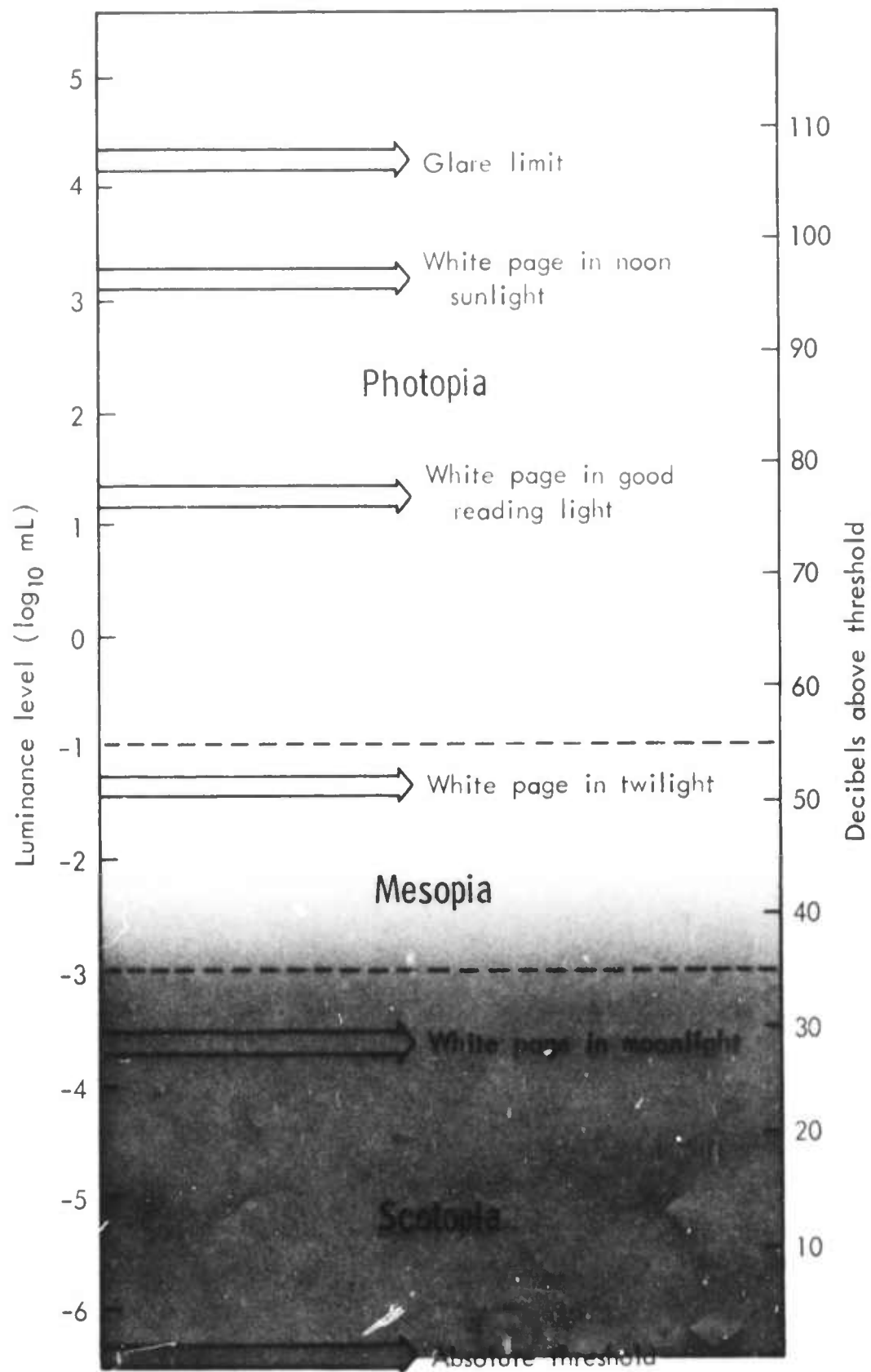


Fig.7—Operative luminance range of the human visual system

The expression $P_\lambda d\lambda$ is the spectral radiant flux corresponding to the radiation between λ and $(\lambda + d\lambda)$. The proportionality constant K is called the "luminous efficiency" and is determined by applying Eq. (1) to the radiation of a full radiator (i.e., a blackbody) at the temperature of solidification of platinum, yielding the value $K = 678.8$ lumens/watt (lm/watt). Thus, luminous flux is that quantity characteristic of radiant flux that expresses the capacity of the latter to produce a brightness sensation in the standard observer.

Now since the CIE has defined both a scotopic function V_λ' and a photopic function V_λ , Eq. (1) may be used to set up two entirely distinct systems of visual photometric quantities, each valid in its own region. The CIE procedure⁽²⁵⁾ is to use common symbols and differentiate the scotopic quantities by a prime, as in V_λ' , above. Specially named scotopic units have been introduced⁽²⁾ in Germany for military problems, but the only fully developed system is the photopic one. Unless otherwise indicated, all luminous quantities herein refer to CIE standard photopic vision.

Finally, specifically embodied in the fundamental definition given in Eq. (1) is the assumption that the luminance of an additive mixture of differing radiations is equal to the sum of the luminances of the components. This rule is known as "Abney's Law" and is photometrically exact by definition; i.e., Eq. (1) is Abney's Law. In addition to satisfying the principles of proportionality between radiant and luminous quantities and of additivity of luminances, the integral nature of Eq. (1) means that there are a nondenumerable infinity of distributions P_λ that give the same value of F . Thus the standard observer may "perceive" as equal in luminance an infinity of radiances differing in spectral composition. The defining relations and units of some radiant and luminous photopic quantities are given in the Appendix.

METAMERS

The brightness response of the CIE Standard Observer has been discussed above without specific reference to any underlying experimental data. The functions were, of course, originally based on experimental

results. The appropriateness of the above approach should become clear in Section IV, which is specifically devoted to the actual retinal sensitivities of human observers. However, before discussing the color response of the CIE Standard Observer, at least a modicum of experimental background is required for the sake of clarity. The pertinent results are those obtained from the so-called matching experiments, in which the human observer is used as a null device, i.e., the observer simply states when two parts of his visual field are indistinguishable.

If a color-normal observer is presented with a two-part visual field and he attempts to adjust one part so that its color matches that of the other, he will in general require three independent adjustments for the controllable element. For example, if he tries to match a single spot of colored light by projecting several spotlights of differing color superposed on a neighboring part of a screen, he generally finds that three lights, each of which has a fixed spectral composition, are required. The same rule applies to rotary mixtures on a sectorized disc and to pigment combinations in paint. Thus, for example, we have the familiar red, yellow, and blue subtractive pigment primaries in grade school paint sets; the magenta, yellow, and cyan subtractive filter primaries of color photography; and the additive spotlight primaries, red, green, and blue.

Now suppose that an observer has obtained a color match by adjusting the three spotlights at his disposal; i.e., he has set up a second stimulus that is equivalent to the first. The observer is said to have made a metameric match,⁽²⁶⁾ and the two stimuli are said to form a metameric pair, or are said to be metamers. Normally, however, the mixture stimulus will not be identical in spectral composition with the comparison stimulus even though the two stimuli appear the same.

It should be noted that metamerism is qualifiable, i.e., it has degrees. Thus in matching a paint patch by using a mixture of paints, the degree of metamerism is likely to be small. But if the paint patch is illuminated by daylight and is matched by illuminating a white card with a mixture of spectral red, green, and blue lights, the degree of metamerism will be very high. Note that it is an oversimplification to state that "the higher the degree of metamerism the 'closer' the

match," since a match is always implied. It is more precise to say that "the degree of metamerism correlates with the precision (repeatability) with which the amounts of the three mixture stimuli may be determined."

Two points should be emphasized:

1. The term "metamers" always implies the result of a "match" by a specific observer (real or defined). Metamers for Observer A are generally not metamers for Observer B.
2. For a given observer, the metamers determined generally depend on the conditions of the matching experiment. If the conditions are sufficiently altered, stimuli formerly accepted as metameric are rejected.

Both of the above are manifestations of a single fact: Metamerism is a psychophysical index of a physiological attribute of a given observer. The "degree" of metamerism is a measure of the "fidelity" of this index. For a given metameric match, the greater the dependence on the visual system of the given observer, and correspondingly, the less the dependence on the conditions of the matching experiment, the greater the degree of metamerism. For example:

1. If two stimuli have a high degree of metamerism for Observer A--
 - (a) He shows high precision (repeatability) in redeterminations of the metamers.
 - (b) The stimuli remain metameric to him over a wide variation in external conditions to which he may adapt.
 - (c) The percentage of the general population that will accept the stimuli as very good metamers is low.
2. If two stimuli have a low degree of metamerism for Observer A--
 - (a) He shows poor precision in successive redeterminations of the metamers.
 - (b) Metamerism is lost to him with relatively small variations in adaptive conditions.
 - (c) The percentage of the general population that will accept the stimuli as tolerable metamers under constant external conditions is high.

THE MATCHING EXPERIMENTS

The matching experiments involve the determination of spectral metamers for various observers. It is important to note the conditions under which these experiments are made, since they differ drastically from those of the daily operation of the visual system. Typical conditions for the matching experiments may be summarized as follows⁽²⁾:

1. The illumination level is sufficiently high that color discrimination is well developed (photopic vision), but not so high as to create glare.
2. The field viewed is completely dematerialized (Maxwellian view); i.e., no object or surface is perceived, and the color appears to float in space.
3. The field viewed is about 2° in apparent diameter, and the subject fixates on the center of the field; the retinal image thus covers the fovea.
4. Two modes of presentation are used: (a) In the step-by-step (or direct-comparison) method, the field viewed is divided into two parts of approximately equal size, the radiant flux from each part being independently controlled with respect to both spectral composition and intensity. The test subject is required only to state when the two parts of the field are indistinguishable. (b) In the flicker method, the two comparison stimuli are alternated in the field at a fixed rate of 10 to 20 times per second. During adjustment of one stimulus, the subject states when the flicker is at a minimum.

The matching experiments generally involve as stimuli two or more monochromatic radiations, plus a radiation continuum called "white." The distribution of spectral intensity of the white stimulus must be specified. For the present discussion it is sufficient to define "white light," W , as being characterized by an equal energy continuous spectrum.

Monochromatic radiation is completely determined by two quantities, the wavelength λ and the spectral radiant flux P_λ . Specification of the visual response of the eye to this radiation also requires two quantities, the wavelength λ and the spectral luminous flux F_λ , where

$$F_{\lambda} = KV_{\lambda}P_{\lambda}. \quad (2)$$

The spectral luminous flux F_{λ} governs the brightness aspect, and the wavelength λ governs the color (hue) aspect of the psychological response.

However, the matching stimuli are normally specified in terms of the luminance L of the viewing field. Spectral luminance L_{λ} of a field is the spectral luminous flux F_{λ} in a given direction (see the Appendix). Spectral luminance will be denoted by $L_i(\lambda_j)$, ($i, j = 1, 2, \dots$). Thus $L_i(\lambda_j)$ denotes L_i luminous units of monochromatic radiation of wavelength λ_j . Note that both L and λ must be given to specify the quantity $L(\lambda)$.

If the viewing field contains a mixture of two or more monochromatic luminances, the visual sensation produced is normally characterized by a single hue. That is, the observer is incapable of distinguishing the hues of the components. For the simplest case of a mixture of two monochromatic luminances $L_1(\lambda_1)$ and $L_2(\lambda_2)$, the result of the matching experiment always falls into one of the following three categories.

Case I. For any monochromatic light of wavelength λ_1 less than about 490 m μ , there exists a corresponding monochromatic light of wavelength λ_2 greater than about 570 m μ such that when the two are mixed in proper proportion the resultant stimulus is psychologically equivalent to the stimulus called "white light." This may be represented by the metameric equation

$$L_1(\lambda_1) + L_2(\lambda_2) \equiv L_w(W). \quad (3)$$

The symbol \equiv should be read as "matches" or "is equivalent in sensation to," to indicate that the experimental origin of Eq. (3) is the matching experiment. A metameric equation such as Eq. (3) represents the psychological equivalence of two stimuli and, hence, is more than a simple algebraic equality. For example, by Abney's Law we obviously have

$$L_1 + L_2 = L_w, \quad (4)$$

and thus Eq. (3) contains Eq. (4), but the reverse does not hold true. Equation (3) states the remarkable experimental fact that the addition of two appropriate monochromatic luminances L_1 and L_2 will match a continuous spectrum of white light (i.e., a colorless sensation). Thus, whereas Abney's Law states that luminances are additive, Eq. (3) states that colors can "neutralize" each other. To illustrate, the "algebra" of Eq. (3) may be represented by Fig. 8, which shows the spectral distributions of the three luminances.

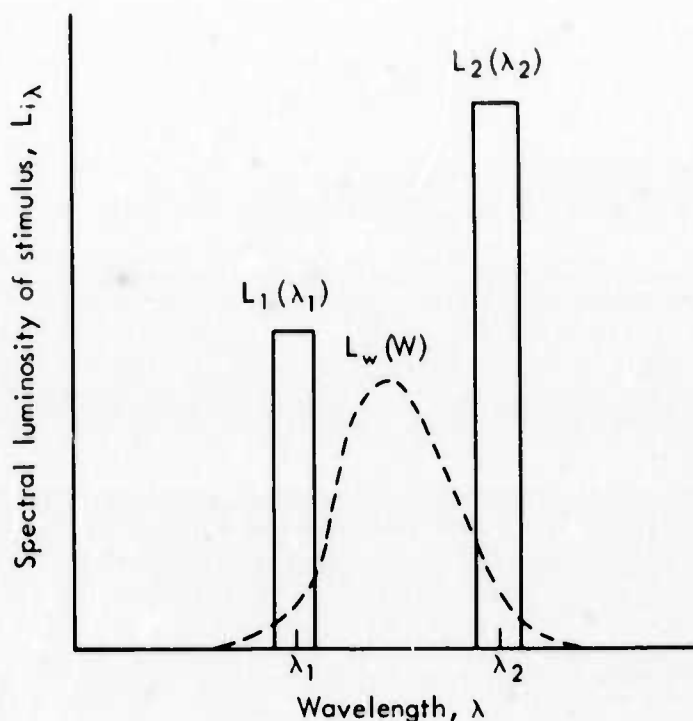


Fig.8— The "algebra" of metamerism

The luminance content of Eq. (3) is thus represented by the equation

$$\int (L_{1\lambda} + L_{2\lambda}) d\lambda = \int L_{w\lambda} d\lambda \rightarrow \int (L_{1\lambda} + L_{2\lambda}) d\lambda = \int L_{w\lambda} d\lambda,$$

whereas the chromatic content is represented by the equation

$$\text{Hue } (\lambda_1) + \text{Hue } (\lambda_2) = \text{Zero hue (white)}.$$

The appropriate pair of wavelengths (λ_1, λ_2) satisfying Eq. (3) are called "complementaries with respect to the specified white W."

If two monochromatic luminances of wavelengths λ_1 and λ_2 not satisfying Eq. (3) are added, the result is one of the following two alternatives.

Case II. In this case, there exists a monochromatic light of wavelength λ_3 that when mixed with white light W will psychologically match the sum of λ_1 and λ_2 , namely,

$$L_1(\lambda_1) + L_2(\lambda_2) \equiv L_w(W) + L_3(\lambda_3). \quad (5)$$

Case III. In this event, there exists a monochromatic light of wavelength λ_4 that when added to the sum of λ_1 and λ_2 will psychologically match white light W, namely,

$$L_1(\lambda_1) + L_2(\lambda_2) + L_4(\lambda_4) \equiv L_w(W). \quad (6)$$

In the case represented by the metamerism equation (6), the sum $L_1(\lambda_1) + L_2(\lambda_2)$ normally produces the psychological sensation known as purple. The appropriate wavelength λ_4 required to satisfy Eq. (6) always lies in the greenish or mid-region of the visible spectrum, whereas λ_1 and λ_2 must lie at opposite ends, i.e., in the orange-red and blue-violet regions.

The metamerism equations (3), (5), and (6) summarize the basic results of the matching experiments. But are these three relations irreducible? Now obviously Eq. (3) is formally contained in Eq. (5), where if L_1 , L_2 , λ_1 , and λ_2 are properly chosen, the resultant L_3 is zero. Also the experimental result represented by Eq. (6) can be put in the form of Eq. (5), namely,

$$L_1(\lambda_1) + L_2(\lambda_2) \equiv L_w(W) + L_3(\lambda_{c3}), \quad (7)$$

where λ_{c3} refers to a "pure purple" light composed of a mixture of two monochromatic lights, e.g., of wavelengths 400 and 700 mμ. The "pure purple" luminance $L_3(\lambda_{c3})$ may in this case be specified by its "purple wavelength" λ_{c3} , where

$$\lambda_{c3} = \frac{L_{400}}{L_{700}}, \quad (8)$$

and by its luminance L_3 , where

$$L_3 = L_{400} + L_{700}. \quad (9)$$

Thus it has been stated⁽²⁾ that Eq. (5) alone is sufficiently general to summarize the matching results. But the economy gained by this step must be balanced against the clarity lost. Finally, it is important to note that Eq. (5) states that it is impossible to match exactly a single monochromatic light with a mixture of lights; i.e., a mixture always appears to contain some white light.

TRIVARIANCE

Equations (5) and (7) are generalized by the experimental fact that the terms in metameric equations obey certain algebraic rules. By their very meaning, all terms in a metameric equation must be positive quantities when the equation represents directly the two halves of the matching field. Both sides of the equation may be increased or decreased by the same term. All terms may be multiplied by the same (positive) factor. The sum of several terms may be replaced by any other sum that is metamERICALLY equivalent to it. Two mixtures, each matching a third, are themselves equivalent. These linear properties, derived from the matching experiments, are known as "Grassman's Laws." Using these laws, we may add a monochromatic luminance $L_5(\lambda_5)$ to both sides of Eq. (5) to get

$$L_1(\lambda_1) + L_2(\lambda_2) + L_5(\lambda_5) \equiv L_w(W) + L_3(\lambda_3) + L_5(\lambda_5). \quad (10)$$

But by Eq. (5) we have

$$L_3(\lambda_3) + L_5(\lambda_5) \equiv aL_w(W) + L_6(\lambda_6), \quad (11)$$

where a is a positive number. Hence by combining Eqs. (10) and (11) we get

$$L_1(\lambda_1) + L_2(\lambda_2) + L_5(\lambda_5) \equiv (a + 1)L_w(W) + L_6(\lambda_6). \quad (12)$$

By repeating this procedure we can reduce the sum of any number of monochromatic luminances, and hence, in the limit, any continuous spectrum, to a mixture of white light and a single colored luminance. The latter would be either a monochromatic luminance or a "pure purple." Since radiation of any complexity may be specified by stating the spectral luminance in terms of N wavelength intervals, we may write our result symbolically as

$$\sum_{i=1}^N L_i(\lambda_i) \equiv L_w + L_\lambda, \quad (13)$$

where N may be as large as desired for accuracy.

Equation (13) demonstrates an extremely interesting fact: Its right-hand side has only three explicit variables (L_w , L_λ , and λ). Now a metamerism equation indicates visual equivalence under the conditions of the matching experiments.

Thus we have the experimentally established result that the "infinity" of variables required to specify complex radiation in physical terms is reduced to three independent variables insofar as a matching stimulus is concerned. This phenomenon is called the "trivariance of color vision." Instead of specifying given radiation by its spectral composition, it would appear sufficient to state an equivalent stimulus by specifying the magnitudes of the appropriate three variables on the right-hand side of Eq. (13).

But here a word of caution must be injected. Trivariance of the right-hand side of Eq. (13) is obtained only by virtue of the prior specification of several "parameters." One such parameter appears explicitly in Eq. (13), namely, the defined white W . The previously specified conditions under which the matching experiments themselves are

conducted constitute "parameters." Implicitly, the definition of luminance itself requires a specification of the luminous efficiency function V_λ for each observer. Furthermore, whereas two lights of identical spectral composition will always appear identical to each observer, two complex lights specified by the same set of three numbers will only appear identical to the observer who measured them. There are wide divergences from one observer to the next. For "trivariance" one must therefore specify as "parameters" the actual color-mixture properties for each observer.

Finally, a word about color sensation: Once a given observer has matched two stimuli of different spectral energy distributions, he will see the two stimuli as being alike (i.e., as being metamers) under a fairly wide variation of viewing conditions. However, the color sensation perceived for the metamers may change considerably. In the matching experiments, the observer has simply stated the existence of a match. He has not described the color sensation he perceives. For example, metamers perceived as orange when originally matched may, after prolonged exposure of the observer to red light, be perceived as greenish-yellow; but they will ordinarily be identical greenish-yellows. The visual system of a given observer cannot be trusted to always yield the same color sensation from a given stimulus.

In summary, we can make the following general statements concerning trivariance of color vision:

1. Under the specified conditions of the matching experiments, the color vision of any given normal observer is trivariant.
2. Metamers are not generally valid from one observer to the next.
3. The matching experiments say nothing concerning the color sensation perceived for a given stimulus.

CHROMATICITY

Colorimetry assumes trivariance of color vision without regard to qualifying conditions. It has created a standard observer that will always give the same "response," to the same stimulus. In view of the preceding discussion, it is necessary to define the color-mixture

properties of this standard observer. The resulting formalism is briefly discussed in this and the following subsections.

If we consider Eqs. (3) through (13) as ordinary algebraic relations, we may combine them into the single colorimetric equation

$$\Lambda = L_1(\lambda_1) + L_2(\lambda_2) + L_3(\lambda_3). \quad (14)$$

Here Λ is any stimulus, monochromatic or continuous, and λ_1 , λ_2 , and λ_3 are three suitably chosen monochromatic stimuli called "primaries." The formerly used equivalent sign (\equiv) has been changed to the equality sign ($=$) in Eq. (14) for a definite reason. Equation (14) does not, in general, express a "match" between the left-hand side and the right-hand side as do metamerism equations. In metamerism equations, all terms are positive additive luminances. In Eq. (14) some of the terms on the right-hand side may be negative; and, in fact, if Λ is a monochromatic light, one component on the right-hand side of Eq. (14) must be negative. The occurrence of negative components in colorimetric equations means that these terms were on the opposite side in the metamerism equations from which the former were derived.

It is obvious from Eq. (14) that the triplet (L_1, L_2, L_3) constitutes a specification of the stimulus Λ . These three numbers are called the "tristimulus values" of Λ based on the primary system $(\lambda_1, \lambda_2, \lambda_3)$. Now the primaries $(\lambda_1, \lambda_2, \lambda_3)$ are not completely arbitrary. From Grassman's Laws we see that the primary system must satisfy two conditions:

1. No one primary is metamerism to a combination of the other two primaries.
2. Some combination of the three primaries will give white light.

These two conditions ensure that all of the metamerism equations summarizing the matching experiments may be algebraically reduced to Eq. (14). There are an infinity of primary systems satisfying these conditions, all interconvertible through linear transformation relations. Hence in the absence of physical reasons for the choice of a standard primary system, criteria of convenience may be used. For this purpose

one may even set up imaginary primaries, imaginary in the sense of not being realizable with a physical light source.

But before the standard primaries are chosen, the color-mixture properties of the standard observer must be specified on the basis of real primaries if the standard is to have any basis in experimental data. Such experimental data were provided by Wright⁽²⁷⁾ and Guild⁽²⁸⁾ in the form of the trichromatic coordinates for the spectrum. Wright used monochromatic primaries 460, 530, and 650 mμ, and his average results for 10 observers are shown in Fig. 9. For a discussion of the meaning and usefulness of trichromatic coordinates (also called "chromaticity coordinates") and other colorimetric variables, one may consult the extensive discussions in LeGrand⁽²⁾ and Judd,⁽²⁶⁾ in addition to Ref. 1. However, the general procedure may be sketched rather simply as follows.

Suppose that for convenience different luminous units (P_1, P_2, P_3) are chosen for the three primaries ($\lambda_1, \lambda_2, \lambda_3$) in Eq. (14). Then Eq. (14) may be written as

$$\Lambda = \Gamma_1(P_1) + \Gamma_2(P_2) + \Gamma_3(P_3), \quad (15)$$

in which the new tristimulus values Γ_i are amounts of the (now unequal) unit quantities (P_i). The trichromatic coordinates are defined as

$$\gamma_i = \frac{\Gamma_i}{\Gamma_1 + \Gamma_2 + \Gamma_3} \quad (i = 1, 2, 3), \quad (16)$$

and hence they satisfy the relation

$$\gamma_1 + \gamma_2 + \gamma_3 = 1. \quad (17)$$

If we now define a new luminous unit for stimulus Λ to be

$$\Lambda = \frac{1}{\Gamma_1 + \Gamma_2 + \Gamma_3}, \quad (18)$$

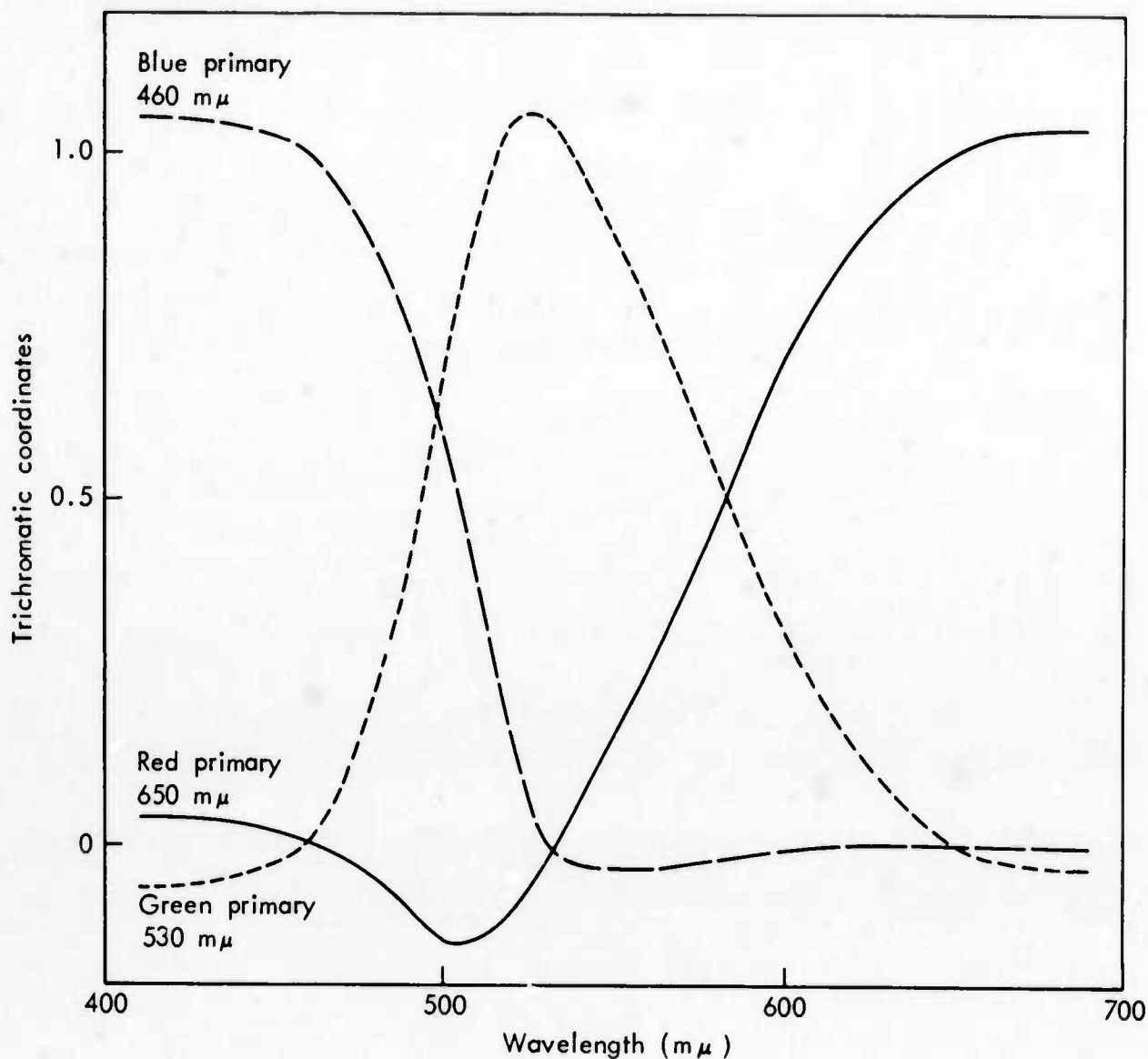


Fig.9—Wright's trichromatic coordinates for the spectrum

then we may write Eq. (15) as

$$\Lambda = \gamma_1(P_1) + \gamma_2(P_2) + \gamma_3(P_3), \quad (19)$$

where the left-hand side now represents one new luminous unit of light Λ . In this rather unnatural system, the trichromatic coordinates

$(\gamma_1, \gamma_2, \gamma_3)$ indicate the percentage that each primary contributes to the total (taken as unity) in "matching" the given stimulus. The coordinates thus serve to emphasize the chromatic aspects of the light Λ as distinct from the luminous aspect. Most colorimetric equations are written in terms of trichromatic (or chromaticity) coordinates in the form of Eq. (19).

The usefulness of the trichromatic coordinates is displayed in the geometric representation of color. In view of Eq. (17) we need only a two-dimensional representation, called a "chromaticity diagram." The spectrum locus for Wright's data of Fig. 9 is presented in a chromaticity diagram in Fig. 10. All physically realizable colors have coordinates falling within the spectrum locus. Note that while the chromaticity of a stimulus is completely specified by its location on the chromaticity diagram, nothing is thereby specified concerning its luminance.

THE CIE CHROMATIC RESPONSE

The aforementioned data of Wright⁽²⁷⁾ and Guild⁽²⁸⁾ were transformed into a new primary system, averaged, and adopted as the color-mixture properties of the Standard Observer by the CIE in 1931. The smoothing of the data ensured that the center of the chromaticity diagram (i.e., $\gamma_1 = \gamma_2 = \gamma_3 = 0.3333$) would represent the equal energy white light source W. This CIE Standard is known as the "RGB System," based on the choice of wavelengths 700, 546.1, and 435.8 m μ for primaries. It is not used for colorimetric calculations for two reasons⁽²⁾:

1. One chromaticity coordinate is always negative for monochromatic radiation.
2. The adopted primaries might acquire some physiological merit that they do not possess. Instead, the CIE adopted the International XYZ System, which is a transformation of the RGB System into a system comprised of three purely imaginary primaries.

The XYZ Primary System has the following features:

1. Since imaginary primaries are used, the new chromaticity coordinates (x, y, z) are always positive for any real stimulus.

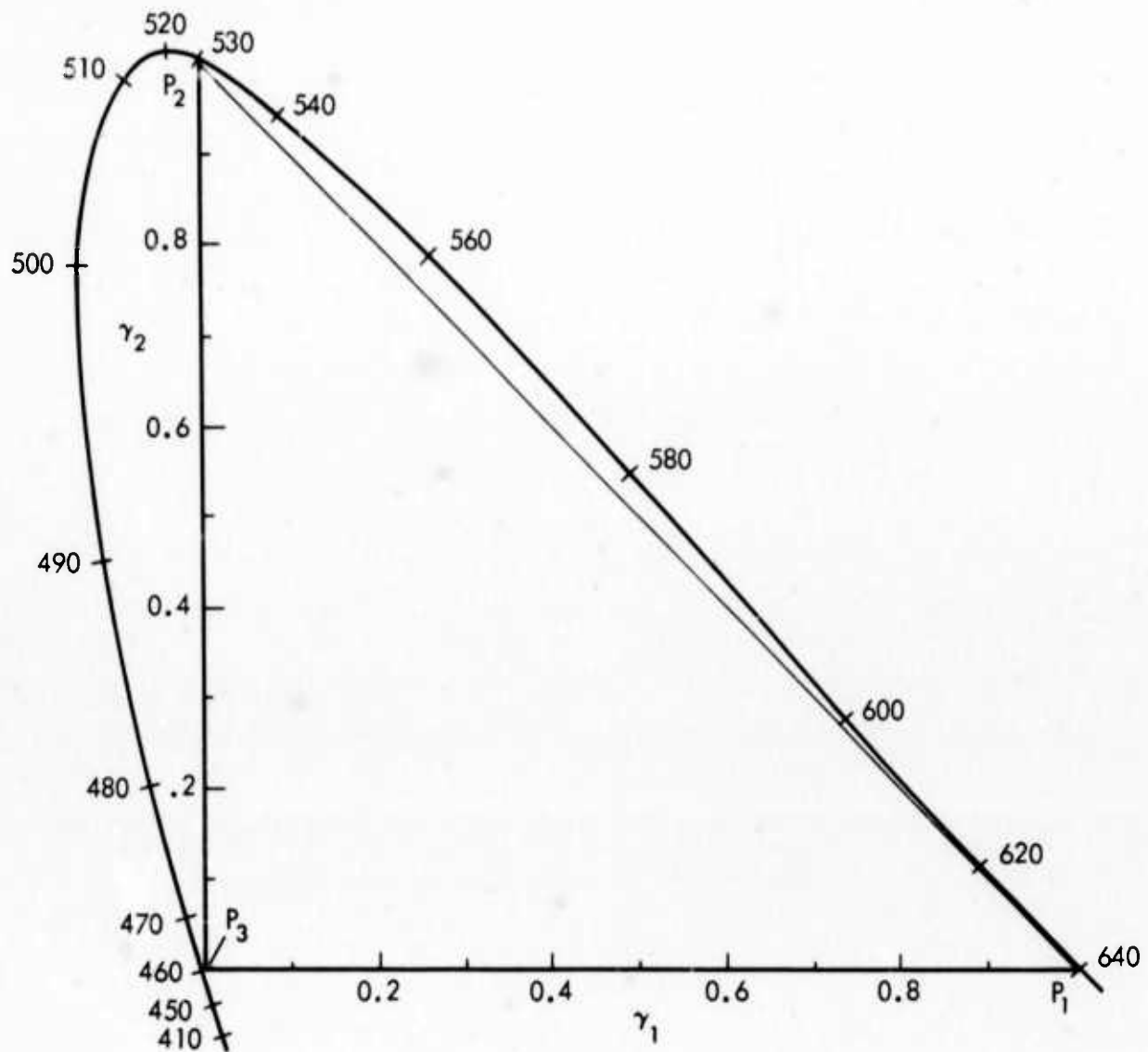


Fig.10—The spectrum locus in a chromaticity diagram for Wright's primaries of 460, 530, and 650 m μ

2. Since luminous units of zero are chosen for primaries X and Z, the luminous unit for Y is thereby unity.
3. Since two luminous units are taken as zero, an additional condition is required to relate the XYZ System to the RGB System. This condition is that the equal energy white W shall be equivalent to equal amounts of the three primaries (X), (Y), and (Z).

The color-mixture properties of the CIE 1931 Standard Observer in terms of the XYZ Primary System are given in Fig. 11, in terms of the so-called distribution coefficients (\bar{x} , \bar{y} , \bar{z}). By definition these are the tristimulus values of the spectrum for the standard observer. Thus the distribution coefficients give the amounts of the three imaginary primaries required to match a unit amount of energy at the given wavelength. The \bar{y} -function is identical with the CIE 1924 Photopic Luminous Efficiency Function V_λ as a result of the choice of the luminous units for the primaries. The tristimulus values for any source of spectral radiant flux P_λ are thus given by

$$\begin{aligned} X &= \int_0^\infty P_\lambda \bar{x}_\lambda d\lambda, \\ Y &= \int_0^\infty P_\lambda \bar{y}_\lambda d\lambda, \\ Z &= \int_0^\infty P_\lambda \bar{z}_\lambda d\lambda. \end{aligned} \tag{20}$$

Since $\bar{y}_\lambda = V_\lambda$, the Y-value is directly proportional to the luminous flux (cf. Eq. (1)). The chromaticity coordinates x, y, and z are obtained from the tristimulus values in analogy with Eq. (16), namely,

$$x = \frac{X}{X + Y + Z}, \tag{21}$$

etc.

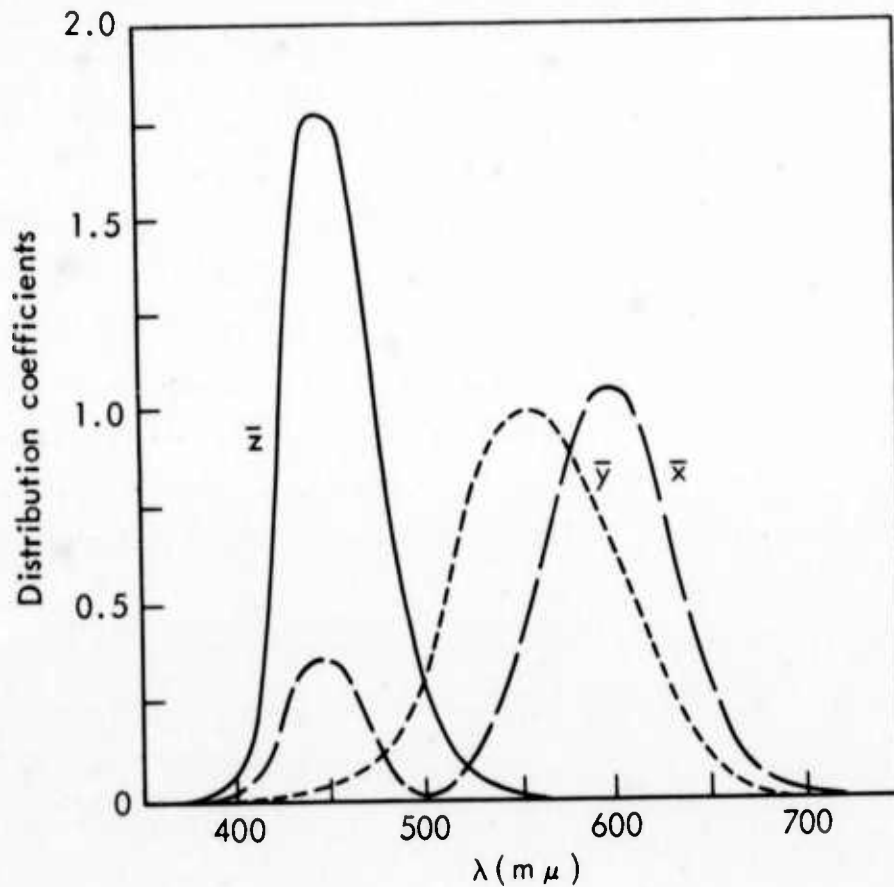


Fig.11 —The distribution coefficients in the XYZ System for the CIE 1931 Standard Observer

The (x, y)-chromaticity diagram for the XYZ System is shown in Fig. 12, where various points on the spectrum locus from 400 to 700 μ are identified by wavelength. The region bounded by this locus and the straight line connecting its termini contains the chromaticities of all physical stimuli. The curved line in the center of the figure is the locus of the chromaticity coordinates for a blackbody source, with the temperatures noted in degrees Kelvin ($^{\circ}\text{K}$).

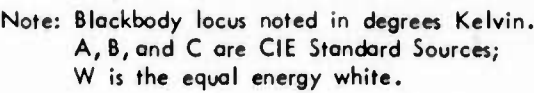


Fig.12 —The chromaticity diagram for the CIE International XYZ System

The chromaticity diagram is a color-mixture diagram in a very real sense. It may be used to determine very simply the chromaticity of the light C resulting from the additive combination of any two lights C' and C'' . The point representing C must be on the straight line joining the points representing the two lights C' and C'' . Further, by Abney's Law and Grassman's Laws, if m units of C' are mixed with n units of C'' , then $(m + n)$ units of C result, and C is located at the point $m/(m + n)$ of the distance from C'' to C' .

The (x, y) -chromaticity diagram has been divided by Kelly⁽²⁹⁾ into 22 colored regions plus white. Kelly also assigned color names to the regions for ease in interpreting chromaticity data. This diagram is given as Fig. 13.

The psychophysical "color" of any given spectral distribution of radiant flux is determinate through colorimetric calculations^{(21), (23)} based on the preceding definitions. Now the only "white" discussed so far is the equal energy spectrum W . To standardize the calculation of the color of the radiant flux reflected from a surface of known spectral reflectance when illuminated with ordinary commercial lighting, the CIE has also defined three artificial light sources that are easy to reproduce and whose spectral energy distributions are precisely known. They are known as Standard Illuminants A, B, and C. Their respective relative energy distributions are shown in Fig. 14 and their chromaticities in Fig. 15. Illuminant A is a gas-filled tungsten lamp operating at a color temperature of 2854°K , and its radiation is characteristic of an ordinary 150-watt incandescent lamp. Illuminants B and C consist of Illuminant A with appropriate defined filters. Illuminant B is similar to the direct light of the sun, while C is similar to an overcast sky or a mixture of sun and blue sky. Under ordinary conditions an observer in an environment illuminated by any of these standard illuminants will perceive as white (or gray) radiation whose spectral distribution is that of the illuminant. Hence the three standard illuminants may be loosely considered as three additional defined "whites" characteristic of practical situations.

Colorimetric calculations with the CIE System always result in three numbers. The customary procedure is to calculate the two

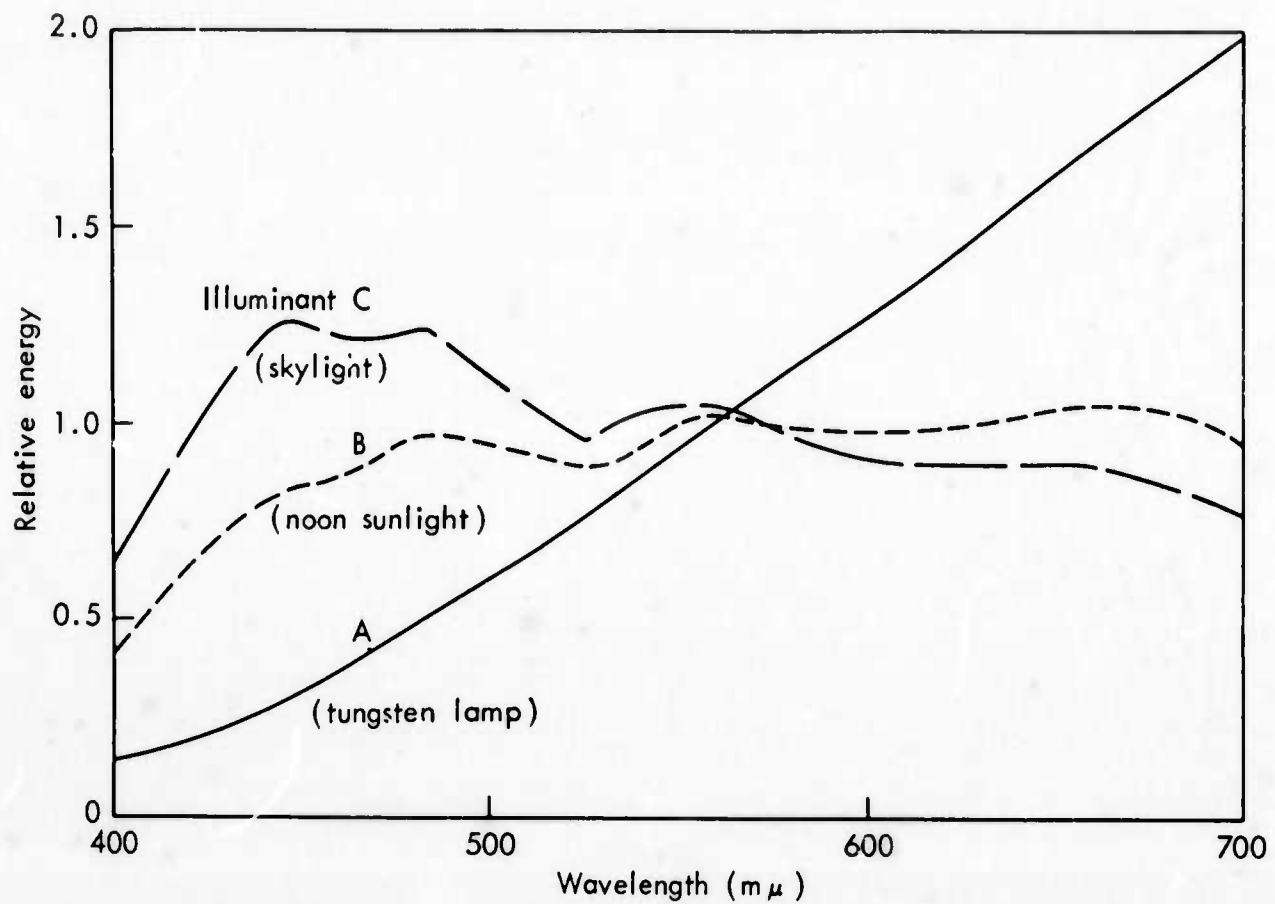


Fig.14—Relative spectral energy distribution of CIE 1931 Standard Illuminants A, B, and C

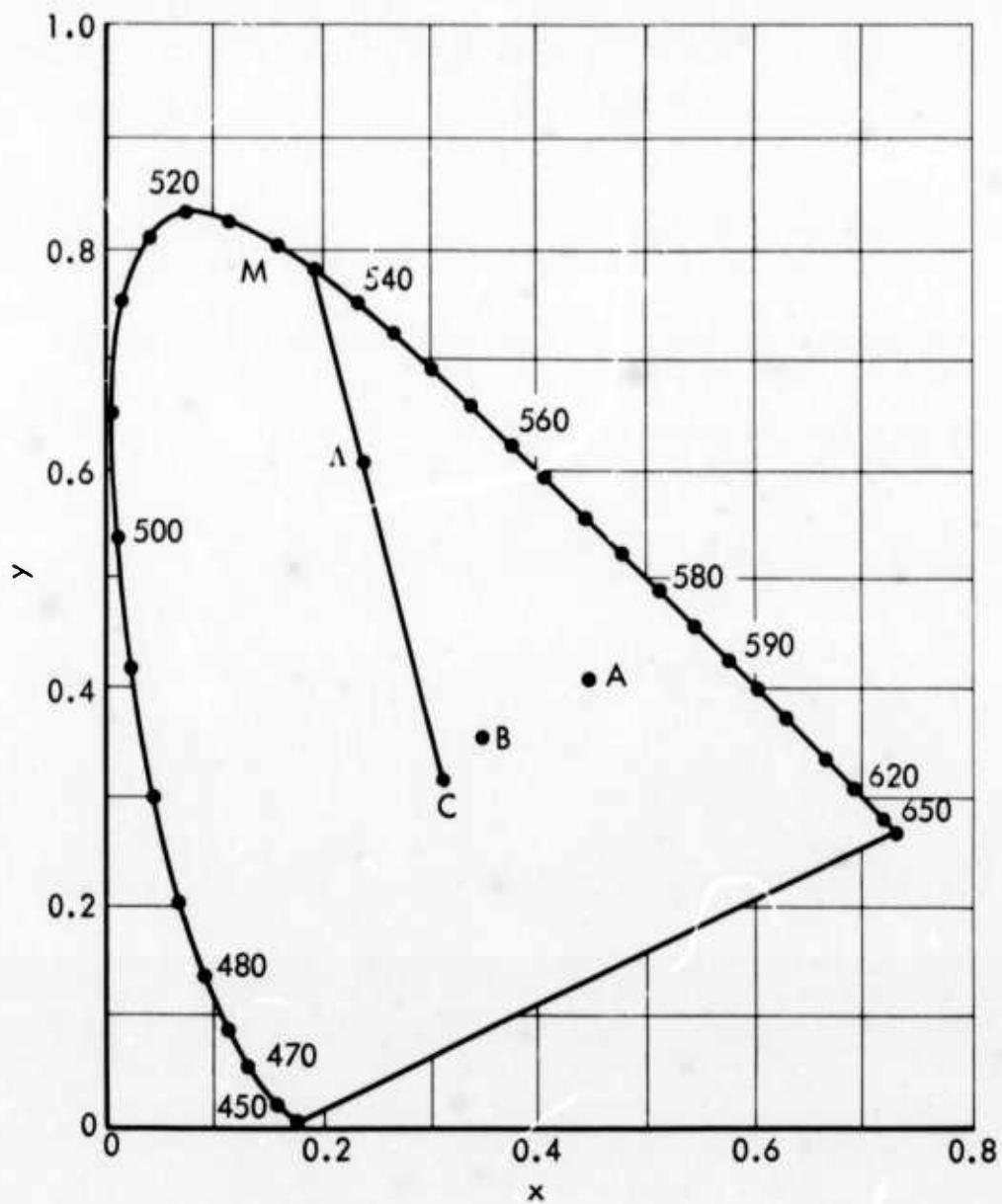


Fig.15 — Specification of a color by dominant wavelength and excitation purity with respect to Standard Illuminant C

trichromatic coordinates x and y , and the tristimulus value Y , which is proportional to the luminance. The result is then plotted as a point on the standard (x, y) -chromaticity diagram, with the value of Y written beside the point. Such a point is shown at Λ in Fig. 15, where Λ may be considered to be the numerical value of Y for the given stimulus. Now the chromaticity of the stimulus may also be specified by its dominant wavelength λ and excitation purity p_e , provided a reference achromatic white point is also specified. Suppose the latter is given as Standard Illuminant C. Then a straight line through C and Λ intersects the spectrum locus on the chromaticity diagram at the dominant wavelength ($\lambda = 535 \text{ m}\mu$ in the example shown in Fig. 15). The excitation purity is defined by

$$p_e = \frac{C\Lambda}{CM} = \frac{x_\Lambda - x_c}{x_m - x_c} \quad (22)$$

Note that the purity of all monochromatic radiation is unity. If point Λ falls below the white point (i.e., in the purples), the complementary dominant wavelength λ_c is used as a specification. The relative luminance Y is thought of as a separate phase (brightness) of a stimulus, whereas the trichromatic coordinates (x, y) , or the dominant wavelength excitation purity (λ, p_e) , specifications are combined into a single concept called "chromaticity."

THE ARTIFICIALITY OF THE CIE RESPONSE

In summary, the above defined basic "visual" properties of the CIE Standard Observer permit the objective specification of the "color" of a given radiant stimulus. It is to this specification that the ordinary psychophysical concept of color refers. This formalism permits evaluation of the physical stimulus in terms of a standard observer capable of "reacting" only to chromaticity and brightness, and to these only in a single manner. It must be emphasized that the specification is restricted to the stimulus itself and implies nothing concerning the perception of a color sensation. Furthermore, surface characteristics,

which may profoundly affect the color sensation perceived by a real observer, have no role in the CIE System. Surface color is inextricably associated with the psychological concept of appearance (see Fig. 5) and includes the idea of a surface viewed in the presence of other surfaces in an illuminated space. Empirical expressions for predicting the appearance of a surface color in certain simple viewing fields were evolved by Judd⁽³⁰⁾ with some success in gross color changes, but such calculations have had limited applicability.*

Finally, in this section an attempt has been made to emphasize the distinction between the experimental facts established by the matching experiments and the formalism of colorimetry derived in part therefrom. This essential artificiality of the CIE Standard Observer has been frequently overlooked. In this connection the CIE has attempted^(25,32,33) to make perfectly clear that it regards its Standard Luminous Efficiency Functions, on which all of photometry and colorimetry are based, as arbitrary wavelength functions adopted for convenience rather than because luminances so evaluated correlate with what a given observer perceives. The same arbitrariness holds true for the distribution coefficients of the CIE Standard Observer, on which CIE chromaticity is based.

*Evans⁽³¹⁾ has recently considered the variables of perceived color and found that at least four independent perceptual variables are necessary to describe the color sensations produced by light sources, illumination, and reflecting and transmitting objects.

IV. RETINAL SENSITIVITY

RETINAL DUALITY

The essential duality of the human retinal receptor mechanism as codified by the CIE in separate scotopic and photopic functions is well established experimentally. Its proof is based primarily on the following phenomena:

1. There are clear anatomical distinctions⁽⁹⁾ between the rod receptors and the cone receptors wherever found in the human retina.
2. Psychophysical measurements⁽²⁾ of the relative luminous efficiency function at high and at low luminance levels result in two distinct functions with their maxima displaying a definite shift in wavelength (the Purkinje Effect).
3. A photopigment, rhodopsin, is found in the rod outer segment, and rhodopsin's absorption properties correlate⁽³⁴⁾ with the spectral sensitivity of scotopic vision.
4. The Purkinje Effect is absent⁽³⁵⁾ in the central fovea, which contains only cones,⁽⁹⁾ in which no rhodopsin has been found.⁽³⁴⁾
5. In visibility threshold measurements, rod-mediated sensations are generally colorless and cone-mediated sensations are generally colored at all wavelengths.⁽³⁶⁾
6. The central fovea, containing only cones, is essentially blind⁽³⁾ in scotopic vision.

These phenomena are discussed more fully in the balance of this study. However, on the basis of the above summary alone, the following conclusions seem straightforward:

1. Rods and cones are different types of receptors.
2. Rods are the principal mediators of dim or scotopic vision.
3. Cones are the principal mediators of bright or photopic vision.
4. Since color vision is well developed only under photopic conditions, cones are the principal mediators of color vision.

In moving from these general, widely accepted statements to the underlying details, one encounters discrepancy and disagreement. It may be said without serious qualification that there is no generally acceptable mechanism known for either the rods or the cones to explain the details of the receptor process that begins when radiation strikes the receptor and ends with the first synapse.

RELATIVE SENSITIVITY

In constructing a mechanism for a retinal receptor, a basic requirement is its spectral sensitivity. Experiment has seemingly satisfied this requirement for the rods. However, insofar as the cones are concerned, the picture is still incomplete. The most extensive available data related to the spectral sensitivities of the retinal receptors are the determinations of the relative luminous efficiencies for different parts of the visible spectrum under various conditions. Such determinations are generally called "relative sensitivities" and constitute a measure of the relative abilities of stimuli of different wavelengths to elicit the sensation of brightness. The knowledge of such functions is fundamental to photometry and colorimetry without regard to the underlying mechanisms, as noted in Section III. But, quite logically, the relative spectral sensitivity of a given region of the retina should display at a minimum some "integrated" effect of the spectral sensitivities of the individual receptors occupying the region. In the present section, an attempt is made to give some indication of the experimental knowledge so displayed. Physiological data are considered in Sections V and VI.

Three experimental methods have found wide use in the determination of relative spectral sensitivity with human subjects. These are as follows:

1. Absolute threshold. The test stimulus is flashed into the field of view for a short time interval, the intensity of succeeding flashes being increased until perception occurs. The reciprocal of the relative energy required for perception is the relative sensitivity. This method is the most straightforward from the standpoint of interpretation.

2. Heterochromatic brightness-match. The test stimulus is matched for brightness with a comparison stimulus in a bipartite field; hue differences in the stimuli are to be discounted. This method is the most difficult for the subject.
3. Flicker photometry. This method depends on the fact that the fusion frequency of repetitive stimuli is a function of their brightness. Hence two stimuli of different hue are alternated in the field at the rate of 12 to 16 flashes per second, and their relative intensities are adjusted until flicker disappears (or is at a minimum). The reciprocals of the relative intensities so determined are quite similar to the relative sensitivities determined by the previous two methods. This method is the easiest to use.

A considerable amount of available experimental data taken with the above methods is summarized in Figs. 17 through 21 and 23 through 25. The experimental conditions used and the number of subjects tested are noted under each figure. The text is concerned principally with general conclusions that can be drawn from the totality of the data rather than from a detailed discussion of each curve.

An immediate general observation is the wide variation among spectral sensitivity curves for various individuals under otherwise constant conditions. This wide variation among individuals for these functions was early noted. Figure 16 shows the spread among foveal relative sensitivity curves for 6 of the 52 observers studied by Gibson and Tyn-dall⁽³⁷⁾ in 1922. At about 620 mμ, the variation in sensitivity is three to one. The differences were believed at that time to be indicative of the diversity existing in the general population. A review of this and other early photopic data was given by Gibson⁽³⁸⁾ in a discussion of the suitability of the CIE 1924 Photopic Standard as a representative "average."

More recent measurements have shown even greater divergences among individuals and have emphasized the effects of factors such as experimental method, size of visual field, image location on retina, luminance level, and degree of adaptation on the measured spectral sensitivity of

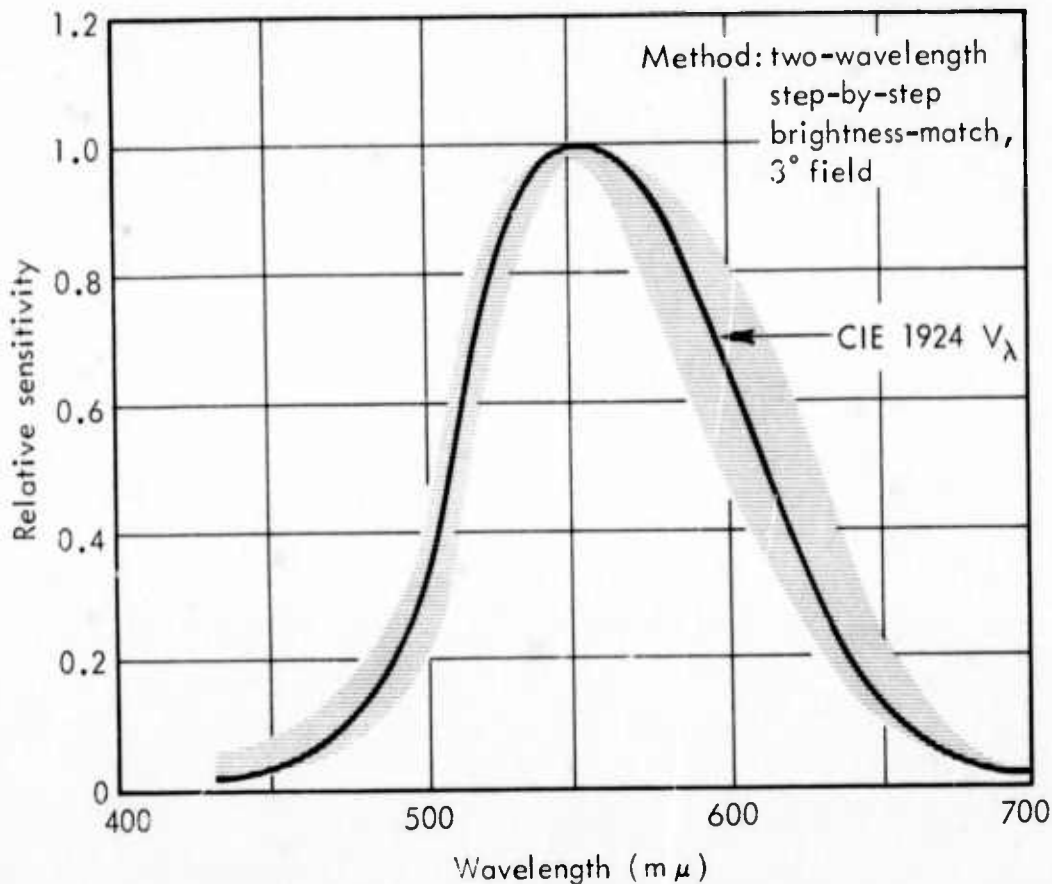


Fig.16 — Total spread of foveal spectral sensitivity for six different persons compared with the CIE 1924 Photopic Standard (Ref.37)

both rods and cones. For discussion purposes, the data are grouped primarily on the basis of luminance level; the periphery is considered in scotopia, and the fovea in photopia.

SCOTOPIC SENSITIVITY

Some of the most careful measurements made of the separate spectral sensitivities of the rod receptors and of the foveal and peripheral cone receptors are those of Wald,⁽³⁶⁾ shown in Fig. 17. The CIE 1951 Scotopic Function V'_λ of Fig. 6 is based essentially on Wald's rod data at 8° above the fovea. Wald's foveal cone (i.e., photopic) data,

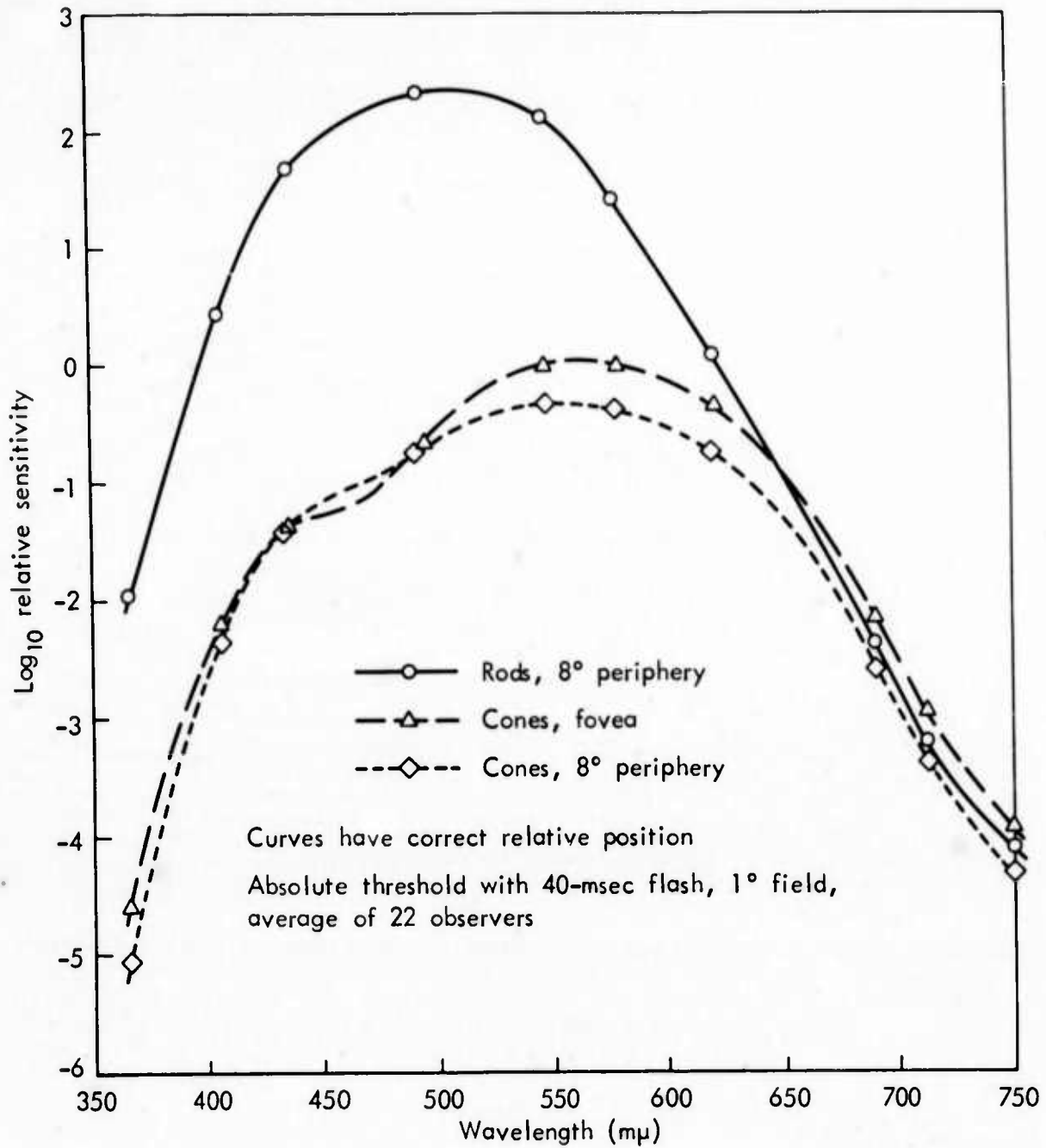


Fig.17—Measurements of receptor spectral sensitivities in fovea and periphery by Wald (adapted from Ref.36)

however, depart widely from the CIE 1924 Photopic Standard V_λ , particularly in the blue-violet. At 436 m μ his data are 2.4 times, and at 405 m μ about 9 times, as high as the CIE Function. Wald himself notes⁽³⁴⁾ that the CIE Photopic Standard in the blue-violet was placed approximately midway between two sets of data that widely disagreed, and he observes that "it is a curious fact that no set of measurements yet made in the human eye conforms at low wavelengths with the International Standard."⁽³⁴⁾ Wald's peripheral cone data in Fig. 17 are the average of 10 of the 22 subjects and were taken with a selective adaptation technique to circumvent the rod activity. They show a maximum sensitivity about 12 m μ below that of foveal cones and a generally lower overall sensitivity. The Purkinje shift is about 57 m μ in Wald's data, from 505-m μ rod maximum to 562-m μ foveal cone maximum. The latter is well above the CIE photopic maximum of 555 m μ .

More recent measurements at much smaller wavelength intervals of the absolute scotopic threshold by Kinney⁽³⁹⁾ are given in Fig. 18. They vividly illustrate the individual variations in absolute thresholds. But the average relative threshold for 4 of Kinney's subjects shows excellent agreement with the CIE 1951 Scotopic Standard, as shown in the summary plot of Fig. 19. Figure 19 clearly illustrates the wide variation at long wavelengths in scotopic measurements found among authors. This variation is usually attributed to cone influence. The reasonableness of this presumption is shown by Wald's results of Fig. 17, wherein the peripheral rods and cones have comparable apparent absolute sensitivities at long wavelengths. Thus the threshold there is probably due to cones alone.

The emergence of cone influence in the periphery is clearly indicated as one moves to luminance levels above scotopic threshold. The changes in individual peripheral spectral sensitivities with increasing intensity levels from scotopic to near photopic conditions have been measured at 10° from the fovea by Walters and Wright⁽³⁵⁾ and Kinney.⁽⁴⁰⁾ These data are shown in Figs. 20 and 21 and clearly display the Purkinje Effect: The peak sensitivity shifts to longer wavelengths with increasing intensity level. But the most extraordinary feature is the appearance, above threshold, of severe irregularities in the sensitivity curves.

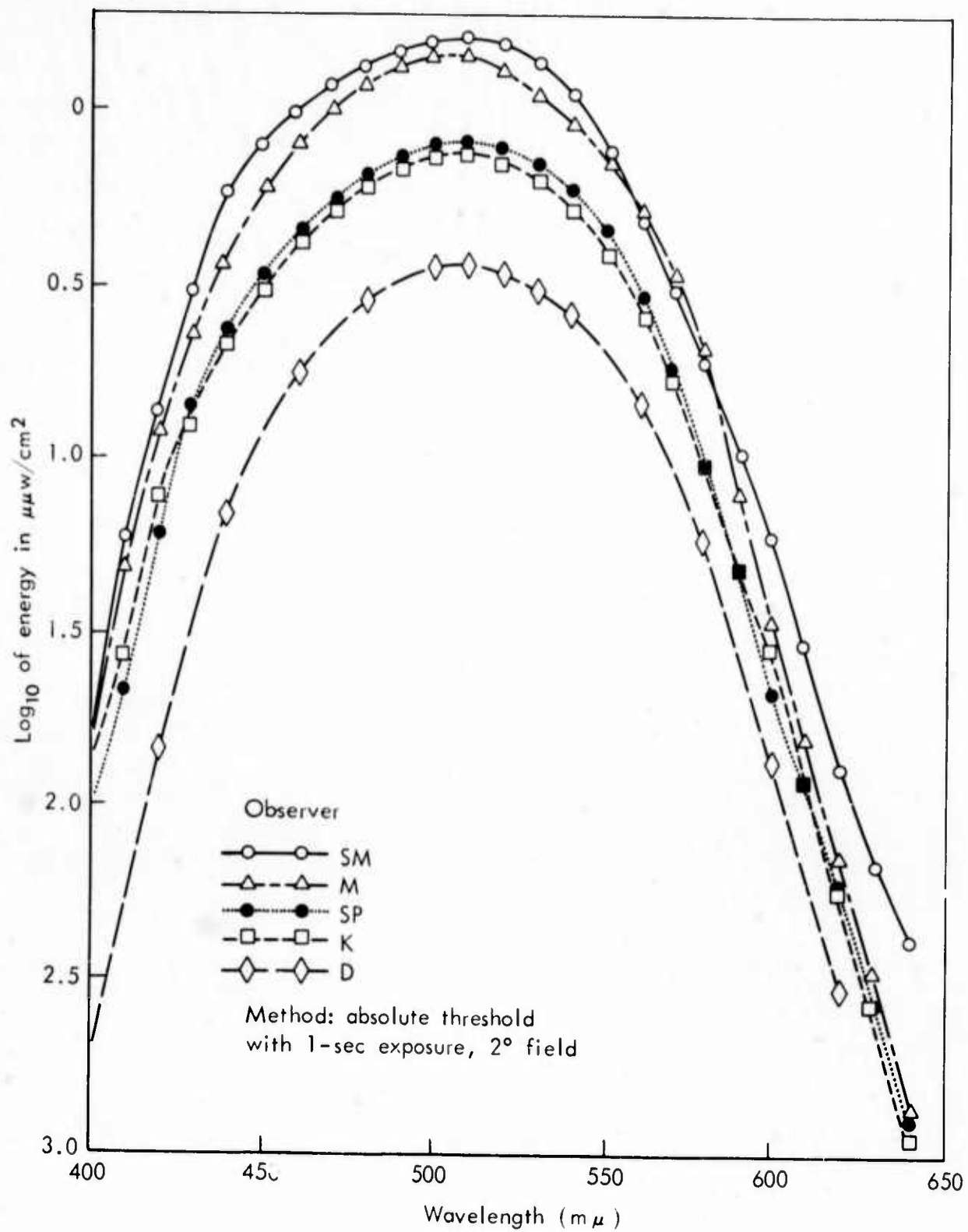


Fig.18—Measurements of scotopic spectral sensitivity at 10° from fovea by Kinney (adapted from Ref. 39)

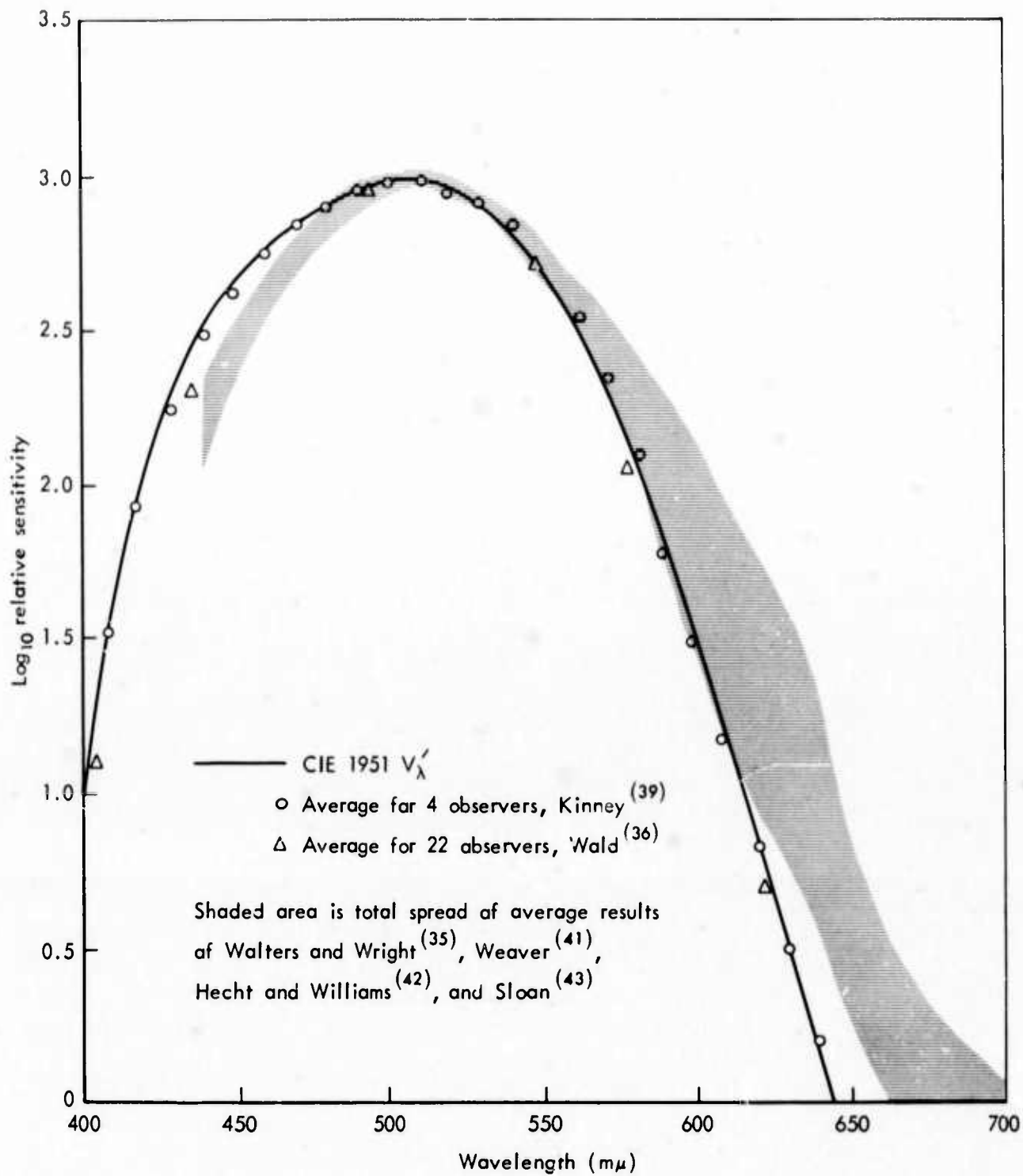


Fig.19 — Comparison of average scotopic spectral sensitivity data from six investigators

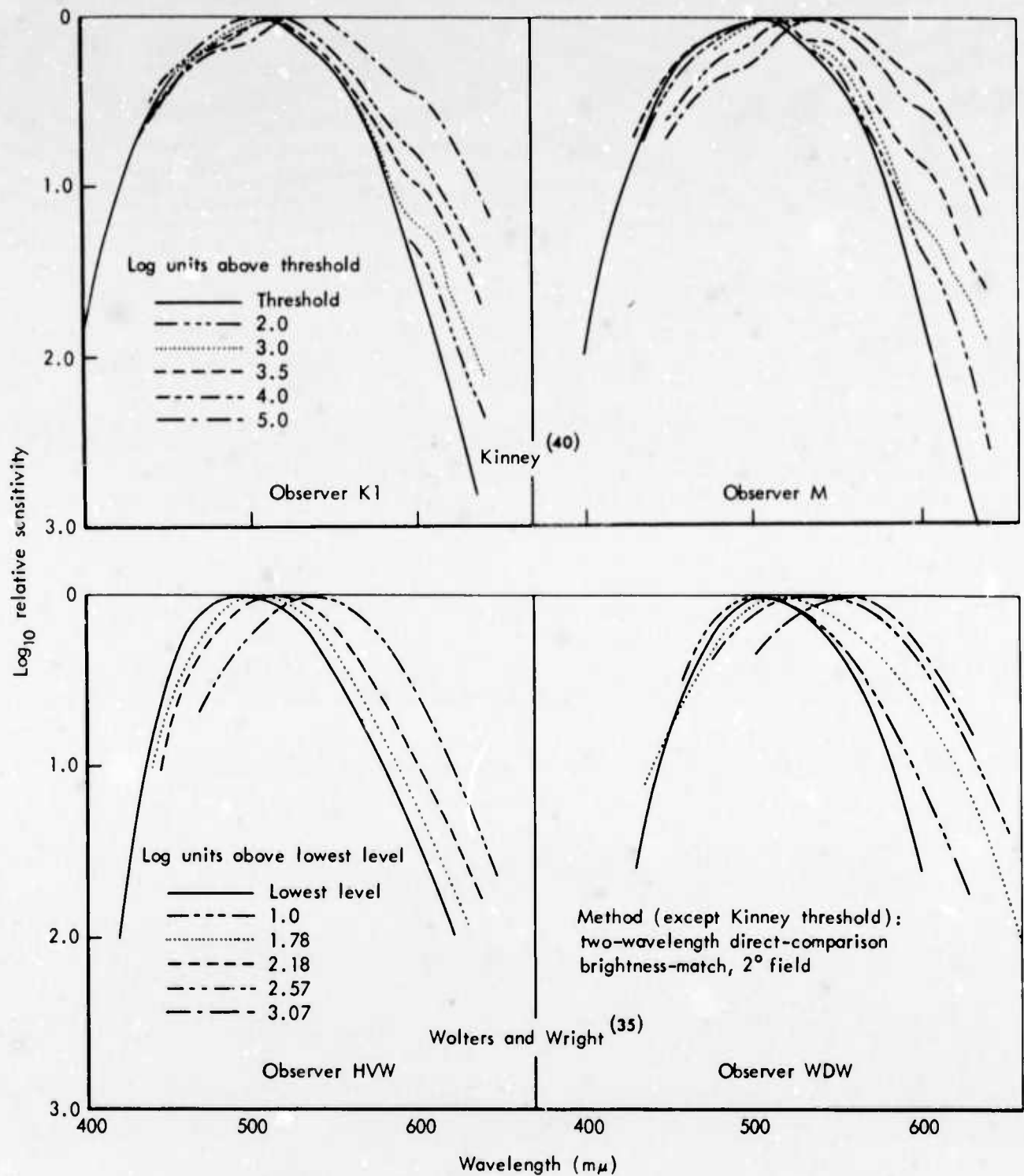


Fig.20—Effect of intensity level on relative spectral sensitivity at 10° from fovea

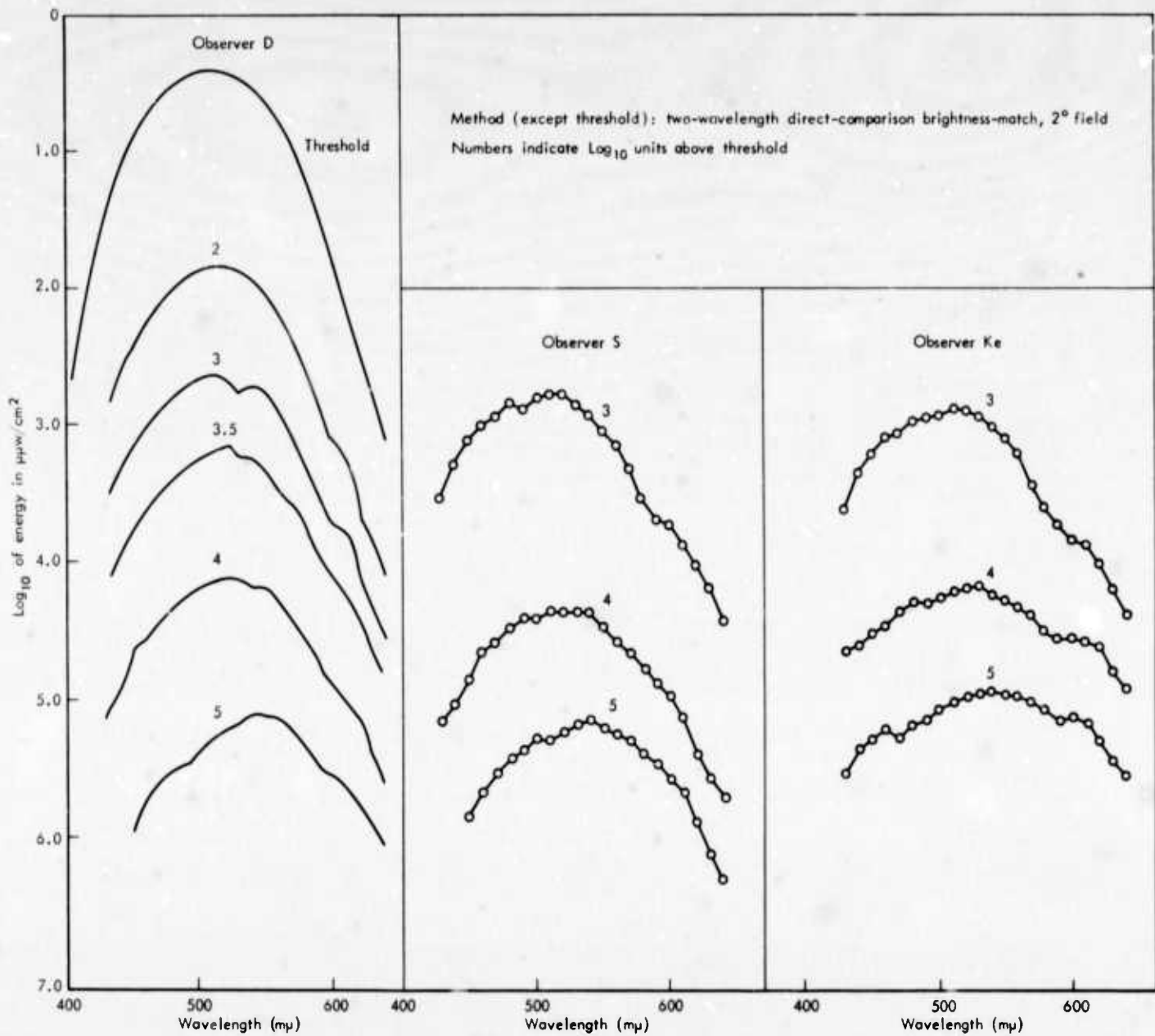


Fig.21—Effect of intensity level on absolute spectral sensitivity at 10° from fovea by Kinney (Ref. 40)

These irregularities are variously called "humps," "peaks," or "secondary maxima." Their factual existence is clearly shown by Kinney's data in Fig. 21. Their essential absence in Walters and Wright's curves of Fig. 20 is explained simply: The authors' results represent smooth fairings with no data points included; further, Walters and Wright displayed an a priori bias toward smooth curves by refraining from reporting irregularities admittedly found in their results. Such an absence of actual data points plagues the literature on color vision.

From a careful perusal of the data presented in Figs. 17 through 21, several experimental facts seem clearly established:

1. The absolute threshold scotopic sensitivity data tend to define unusually smooth experimental curves, particularly at wavelengths below about 560 m μ . No evidence of secondary peaks (or humps) is apparent.
2. While absolute threshold sensitivities show significant variation among individuals, this variation is of a "simple" character. The relative threshold sensitivity is sensibly constant among individuals, again particularly so below about 560 m μ .
3. At intensity levels up to about 2.5 times threshold ($0.4 \log_{10}$ units), the variation in relative scotopic sensitivity with intensity is small or absent below about 560 m μ . Above 560 m μ , variation in relative scotopic sensitivity with intensity is appreciable and continuous from threshold up; evidence of humps appears quickly in this spectral region as one moves above threshold.
4. As the intensity level is increased through scotopia and into mesopia, definite secondary maxima (humps) appear in individual sensitivity curves. These humps appear first in the region above 580 m μ and are clearly defined at intensity levels 100 times threshold. At levels from 10^2 to 10^3 times threshold, humps also appear in the mid-region of the spectrum, from 480 to 580 m μ . Humps below 480 m μ appear at levels 10^4 to 10^5 times threshold.
5. While all individuals show the same general trends mentioned

in point 4, above, there are considerable variations among individuals regarding absolute sensitivity, number of and spectral location of peaks, and magnitude of peaks. The only apparent correlation in peak location among individuals is in the development of the first peak at the long wavelength end; the 5 Kinney observers show definite similarities here up to intensity levels of 10^3 times threshold.

A number of conclusions may be drawn from the preceding results. At scotopic threshold, peripheral sensitivity is due to a single receptor mechanism, the rods, except at long wavelengths where the rod and cone mechanisms apparently are not separable. A single type of receptor with a single relative sensitivity curve is involved in the rod mechanism. The appearance of humps in scotopic sensitivity curves indicates involvement of the cone receptor mechanism. The number, location, and behavior of these humps with intensity changes indicate that the cone mechanism is complex; no separation into components is immediately apparent.

Finally, the mesopic range from scotopic to photopic levels is very broad, with no clear boundaries at either end. At long wavelengths, particularly from about 600 mμ on, there is a continuous change in peripheral sensitivity from absolute threshold on up, due to cone participation. Hence no accurate peripheral sensitivity value can be given without reference to the appropriate intensity level. So far as scotopic threshold is concerned, the relative sensitivity function of the CIE 1951 Scotopic Standard seems to be a very adequate and experimentally meaningful average for real observers. However, the results indicate that no average relative sensitivity curves of comparable meaning are possible significantly above threshold. Hence essentially arbitrary defined standards must be provided in this range for photometric and colorimetric purposes. Although mesopic functions have been proposed by Grigorovici and Aricescu-Savopol,⁽⁴⁴⁾ as shown in Fig. 22, it would seem equally reasonable and much more convenient to define analytic functions as standards in this range.

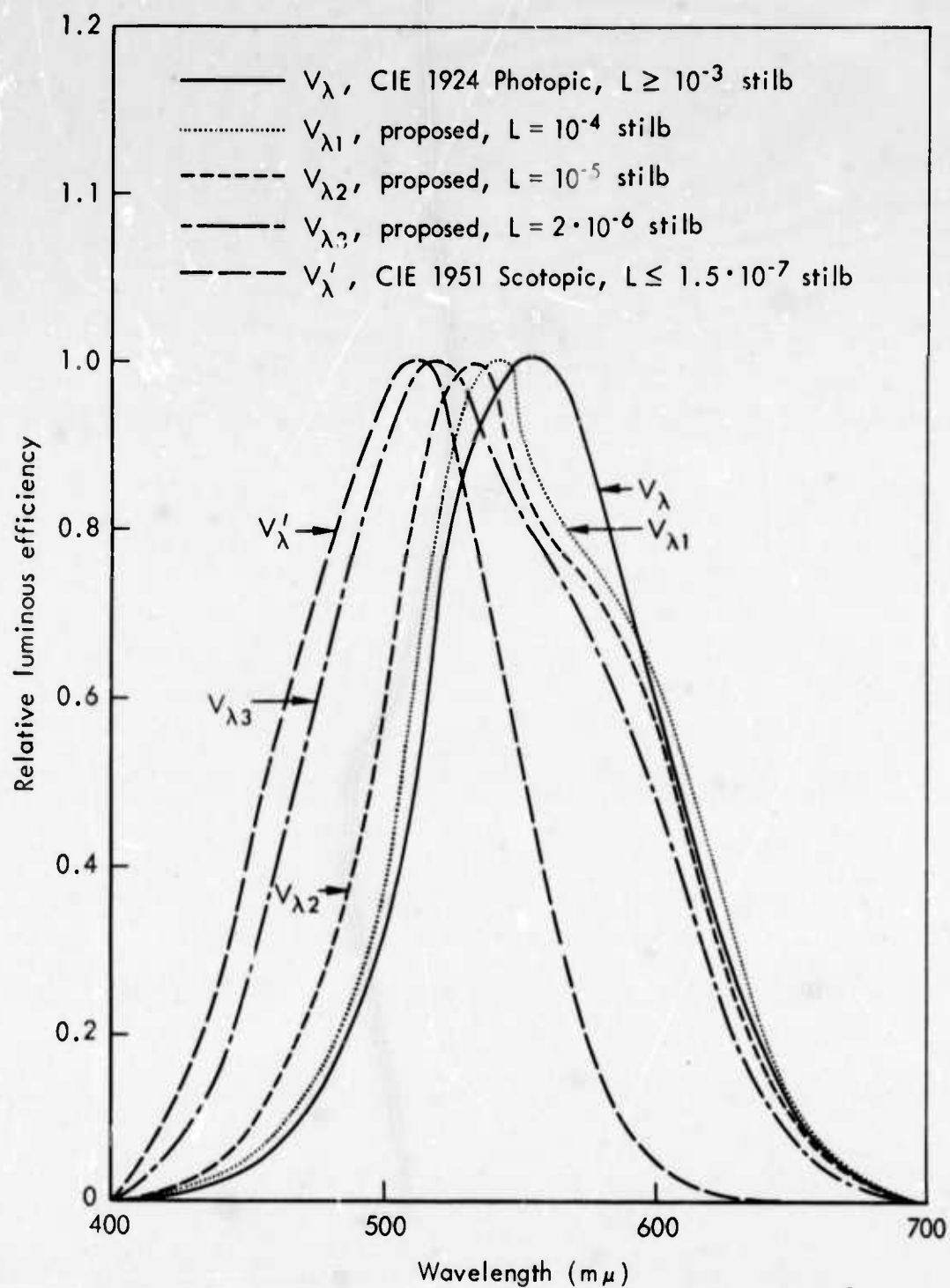


Fig.22—Proposed relative luminous efficiency functions for the mesopic range (Ref. 44)

PHOTOPIC SENSITIVITY

It was concluded in the preceding subsection that scotopic (rod) spectral sensitivity is due to a single receptor mechanism. The photochemical basis of this mechanism will be further discussed in Section V. Of course, color vision is absent in scotopia, and we now turn to a consideration of retinal sensitivity where the cones are the primary mediators, i.e., photopia.

We shall concentrate almost exclusively on the fovea, where the least influence of the rod mechanism is to be expected.⁽⁹⁾ The adjectives "photopic," "foveal," and "cone" are used roughly interchangeably to indicate a foveal cone response whether at threshold or at high luminance levels. This should not be construed as implying that the author feels the periphery to be unimportant, or that the peripheral cones do or don't differ from their foveal brothers. The reasons for the present approach are simple. The general purpose of this study is to seek established experimental facts pertinent to human color vision. Color vision has its highest development in a foveal region approximately 2° in extent. Furthermore, for this region alone, one encounters in the literature an abundance of experimental data taken under an almost equal diversity of experimental conditions. Since variations among individuals are much larger here than in scotopia, this multiplicity of experimental procedures renders the photopic data orders of magnitude more difficult to interpret. We simplify our task considerably by limiting our consideration to data taken with central fields of 2° or less.

The reasons for the widespread concern with foveal relative sensitivity are twofold. First, it seems reasonable to assume that there is a fundamental connection between the foveal relative sensitivity curve of an observer and his mechanism for color vision. Second, such curves are relatively easy to obtain experimentally. The problem, of course, is the interpretation of these data with regard to any inherent information about underlying receptor mechanisms. The literature abounds in such interpretations. But from the present writer's purview, many of these interpretations suffer from one or both of two fundamental

defects: (1) a preconception of what such data must contain and (2) the promulgation as general experimental fact of a result based on a very few (or even single) subjects. On the other hand, if one attempts to reinterpret the available data with a conscious avoidance of these errors, one encounters many interesting investigations whose results are useless because of the mode of reporting. To an experimentalist, data mean data points. Curves determined by spaced data points with such data points unreported are experimentally meaningless. Furthermore, data points without some index of repeatability and without explicit specification of all parameters that may have a bearing on their value are only conditionally acceptable. Finally, in view of variability, average values for two or more subjects must be avoided.

In considering foveal relative sensitivity data we shall therefore concentrate on curves defined by reported data points for individual subjects. A collection of such determinations is presented as Figs. 23 through 25, grouped according to the method used. Figure 23 shows individual foveal threshold curves for a range of field sizes from 3' to 2°. The flash duration and data source are indicated for each curve. Figure 24 presents individual foveal luminosities as determined by direct-comparison brightness-matches for field sizes ranging from 1.5' to 2°. The data source and retinal illuminance in trolands* are given for each curve. Similarly, Fig. 25 presents individual foveal luminosities determined by the flicker method for field sizes from 50' to 2°. The data in Figs. 23 through 25 are taken from Refs. 40 and 45-63.

Figures 23 through 25 lend convincing evidence of the irregular shape of individual foveal sensitivity curves, which seemingly always display "humps" (or "peaks" or "dips") at one or more regions of the spectrum. While early noted by Sloan,⁽⁴³⁾ these irregularities were not uniformly reported even when found,⁽³⁵⁾ apparently because the investigators considered them artifacts. Ten successive determinations of threshold with 1° field and a white surround (10 mL) for one subject

* A troland is the retinal illuminance produced by viewing a surface of 1 candela/meter² (cd/m²) luminance through a pupil of 1 square millimeter (mm²) area.

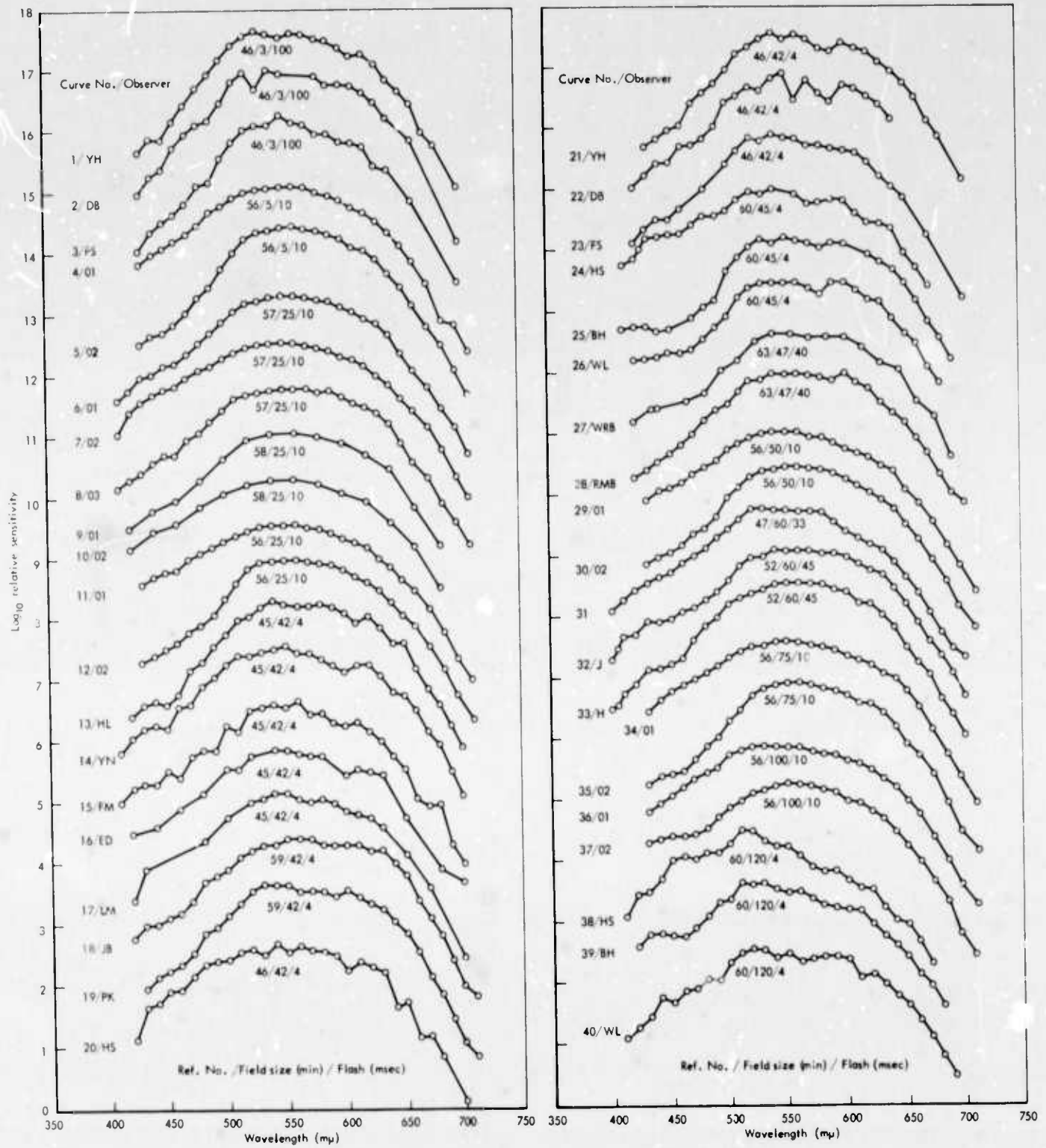


Fig.23 —Determinations of individual foveal thresholds

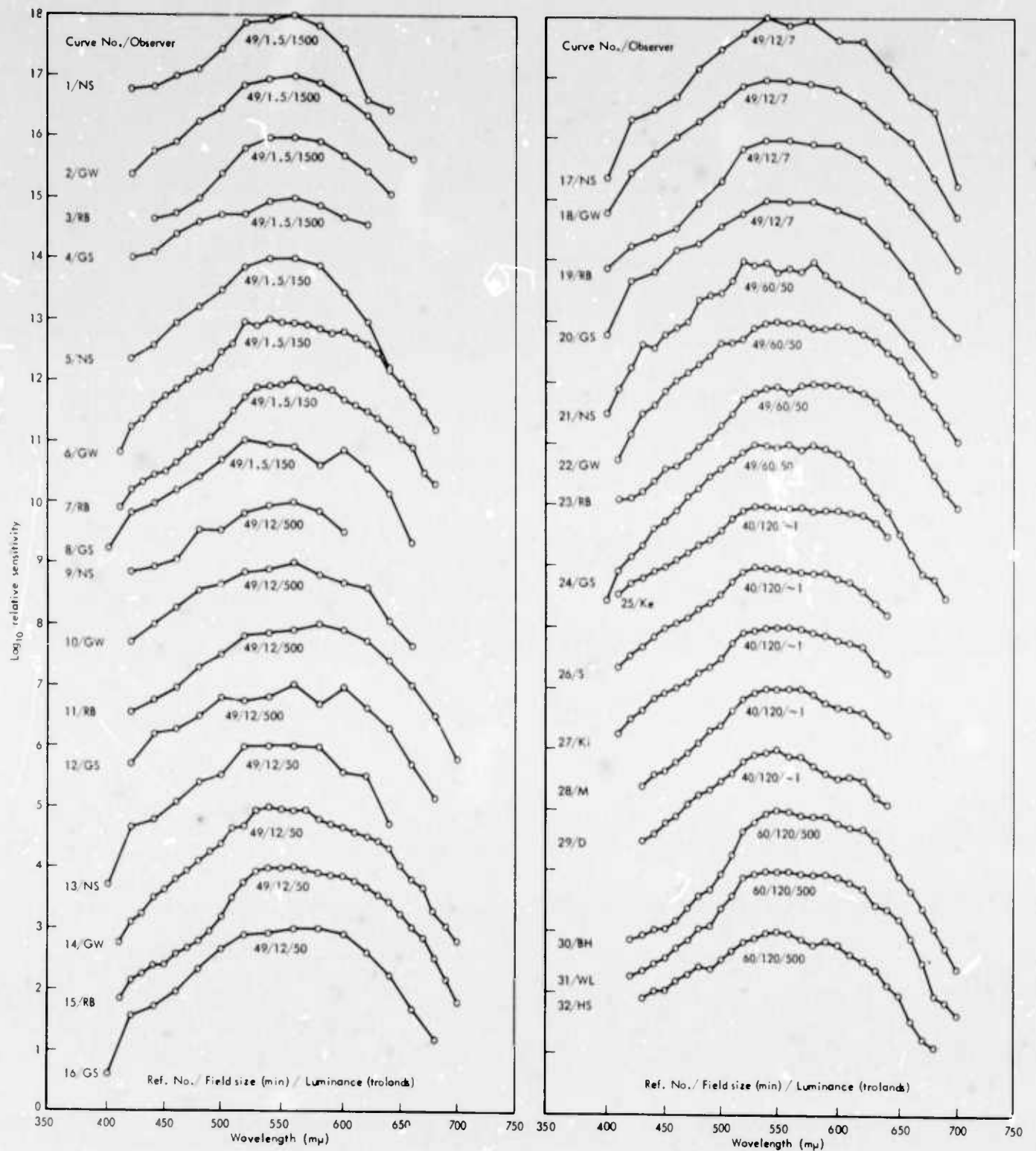


Fig.24—Individual foveal spectral sensitivity determinations by the direct-comparison brightness-match method

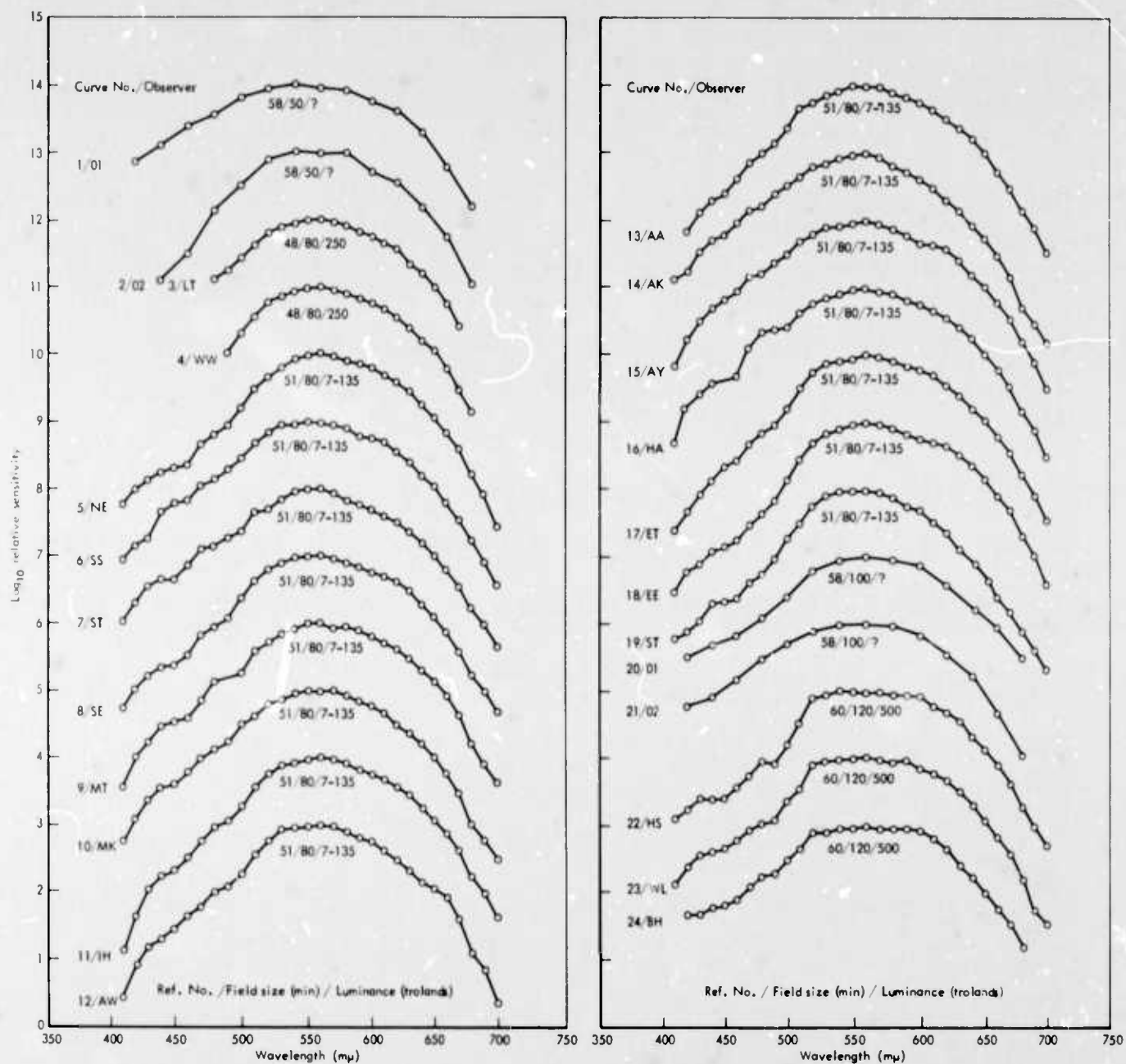


Fig.25—Individual foveal spectral sensitivity determinations by the flicker method

by Hurvich and Jameson⁽⁵²⁾ showed marked changes from run to run in number, spectral location, and size of the irregularities. Sequential data were not given for threshold curves with dark surround, but the data points for a large number of such determinations with a single subject show considerable scatter, as indicated in Fig. 26. Such a figure casts doubts about the validity of fine detail in reported luminosity curves.

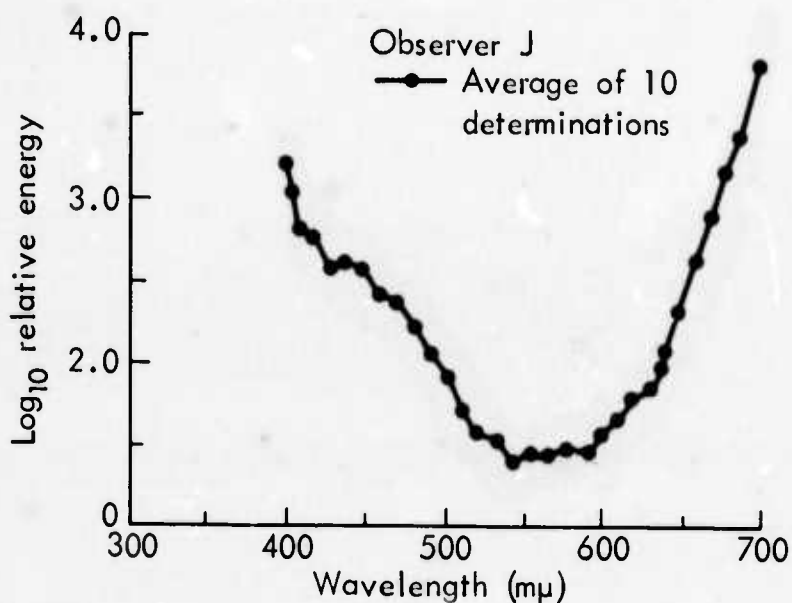


Fig.26—Average and total spread of foveal threshold data for one subject by Hurvich and Jameson (Ref.52)

Nevertheless, several investigators have attempted to verify the reality of certain dips and humps in individual V_λ curves. Specific irregularities for given subjects have been reported^(48,60,62) to persist from one test session to the next. The experimental uncertainty in V_λ determinations is frequently reported as about one-half the magnitude of the major irregularities.* And despite the above-mentioned nonrepeatability of fine detail, Hurvich and Jameson concluded that

*See Refs. 40, 49, 51, 56, and 60 on pages 167 and 168.

there is statistical significance in the major features.^{(52), (64)} But even if the irregularities are accepted as real, the varying experimental procedures and limited number of subjects in each investigation demand caution in their interpretation.

We shall next consider individual foveal spectral sensitivities (denoted as V_λ , for convenience) with regard to (1) the significant irregularities in V_λ , (2) the principal experimental factors affecting the shape of V_λ , and (3) the implications of V_λ determinations relative to the brightness and color discrimination mechanisms.

Irregularities in V_λ

An overview of the curves in Figs. 23 through 25 and their attendant references shows that four major irregularities in individual V_λ determinations have been most frequently reported: (1) A hump (higher sensitivity) in the red is variously reported from 600 to 620 m μ in Refs. 35, 40, 43, 45-47, 49, and 61. (2) A hump in the violet below 450 m μ and one in the blue at about 480 m μ are reported in Refs. 40, 45-49, 51, and 52. A dip (lower sensitivity) at 450 m μ is also reported in Refs. 40, 45, 46, 49, and 59. (3) A double peak in the maximum of V_λ in the green, at 500 to 560 m μ , is reported in Refs. 46, 49, 50, 52, and 55. (4) A sharp dip in the yellow at about 570 m μ , with or without a peak at about 590 m μ , is reported in Refs. 46, 49, 51, 52, 60-62, 64, and 65.

But such "correlations" as the above cannot be assumed indicative of the general population, as the variation in Figs. 23 through 25 shows. For example, with reference to point (4), above, the specific absence of the yellow dip under "expected" conditions is noted in Refs. 45, 51, 52, 57, and 59. Neither can the approximate spectral locations of the above features be given special significance. A careful perusal of the data in Figs. 23 through 25 will permit one, by selection, to establish a "correlation" among several subjects of a "hump" located at almost any desired wavelength. This is verified in Fig. 27, which is a frequency-of-occurrence analysis for the peaks and dips in Figs. 23 through 25. Only the dip at about 580 m μ in the yellow shows any correlation, and this is tenuous.

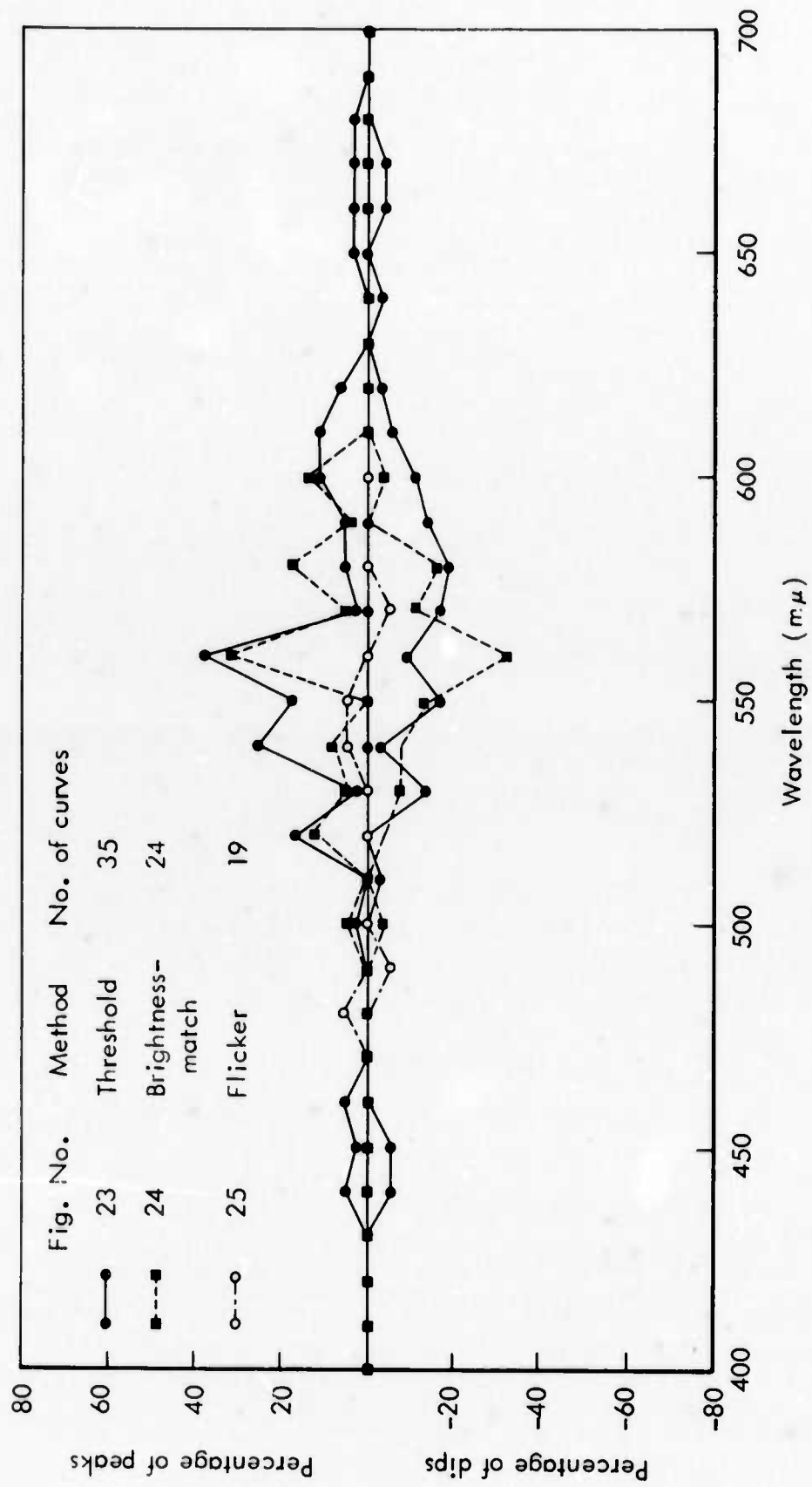


Fig.27 —Frequency of occurrence of peaks and dips in Figs. 23-25

Attempts to relate specific irregularities in V_λ to underlying visual mechanisms are fraught with danger. This is well illustrated by the "saga of the yellow dip." A pronounced dip in spectral luminosity at about 580 m μ has been reported by Sperling and Lewis⁽⁶⁰⁾ as highly correlated among three subjects. Their results indicate that the dip was more severe in threshold data (Fig. 23, curves 24-26 and 38-40) than in photopic flicker and brightness-match data (Fig. 24, curves 30-32, and Fig. 25, curves 22-24, respectively). They concluded from this result that at absolute threshold there is a "failure of brightness summation of the red and green response mechanisms."

This conclusion, however, seems unwarranted. The results of other investigations indicate to the present writer that the variability in the yellow dip found by Sperling and Lewis is a result of their experimental procedures. For example, Heath⁽⁶¹⁾ used several narrow spectral bands in his photopic flicker work and found a pronounced dip at about 570 m μ . Sperling and Lewis⁽⁶⁰⁾ used a "white standard" for their flicker data, and the yellow dip found at threshold was removed. Their brightness-match method used a single monochromatic stimulus (520 m μ) as a standard, and here the yellow dip was fully retained in photopia for one subject (Fig. 24, curve 32). Stiles has reported⁽⁵³⁾ that matching inhomogeneities in the central field were visible in his 2° brightness-match tests only with highly saturated stimuli or stimuli composed of an additive mixture of monochromatic lights. The inhomogeneities were "hardly noticeable" when matching a standard "bluish white."

Wright⁽⁶⁵⁾ has investigated the nature of the yellow dip at low intensities using 20' test objects with a 630-m μ standard and the brightness-match method. He found that the dip did not become more pronounced as the intensity level was progressively decreased toward threshold. The Sperling-Lewis partial failure of summation hypothesis would predict that it should. Rather, Wright found that higher intensities and larger fields removed the dip. Furthermore, Bedford and Wyszecki⁽⁴⁹⁾ found a pronounced dip at 580 m μ in one of their subjects for field-size/luminance-level conditions of 1.5'/150 trolands and 12'/500 trolands (Fig. 24, curves 8 and 12, respectively). But this dip was absent under the 12'/50 troland, 12'/7 troland, and 1°/50 troland

conditions (Fig. 21, curves 16, 20, and 24, respectively). The intensity-level/yellow-dip correlation of their 12' results is thus opposite to that expected under the Sperling-Lewis hypothesis.

Finally, one group of V_λ curves in Figs. 23 through 25 is striking for its absence of severe irregularities. These curves are the determinations by LaDriere⁽⁵⁶⁾ and Collins^{(57),(58)} (see Fig. 23, curves 4-12, 29, 30, and 34-37). These data were all taken at Fordham University. One notes that the severe irregularities displayed by other V_λ curves in Figs. 23 through 25 are about 0.5 log unit, whereas no Fordham determination exhibits a dip or peak greater than 0.2 log unit. Extensive analysis of the Fordham procedures by Dillon and Zegers⁽⁶⁶⁾ and Jackson⁽⁶⁷⁾ places the uncertainty in the Fordham V_λ data at 0.1 to 0.2 log unit, which is of the order of the irregularities in their data. Hence Zegers⁽⁶⁸⁾ attaches no statistical significance to the severe peaks/dips reported in the literature.

Clearly, the results of the above investigations, taken as a whole, imply the following conclusions concerning irregularities in individual foveal spectral sensitivity determinations:

1. Although severe irregularities are probably artifacts, slight irregularities remain, which are most likely anatomically based. But none of the latter have yet satisfied reasonable criteria for statistical significance when compared with experimental uncertainties.
2. If a particular irregularity is present in a given subject, its explicit appearance in a V_λ determination is enhanced by the use of monochromatic stimuli and small fields.

In view of point 2, above, the details of any irregularities would also be expected to depend on the location of the stimulus on the fovea, in addition to the adaptation state^(62,64,69) of a given subject.

Factors Affecting V_λ

The preponderance of experimental data to date indicates that it would be easier to list experimental factors that do not affect V_λ

determinations. There seems to be conclusive evidence that the three principal methods in use--threshold, brightness-match, and flicker photometry--do not give identical results. The frequency used in the flicker method seems important independent of the luminance level,^{(58),(61)} yet it cannot be so simply isolated. Both slow pretest adaptation and rapid local adaptation are important.* These results imply that the duration of the test stimulus affects the resulting V_λ curve. Adaptation effects with the rather slow brightness-match method could only be isolated by extrapolation, a dangerous practice in visual research. In short, the details of the experimental procedure, the state of adaptation, the location on the fovea, the field size, and the intensity level all affect the resulting V_λ curve. We shall give further consideration to only two of these factors, field size and intensity level. Available data show that the two are interrelated.

There are two important and related questions here: (1) To what extent and over what area is there a rod influence on the foveal V_λ ? (2) To what extent is the center of the fovea dichromatic (tritanopic)? With regard to the former, Sperling and Lewis⁽⁶⁰⁾ concluded from the higher sensitivities in the blue for their 2° threshold data as compared with their $45'$ data (see Fig. 23, curves 38-40 and 24-26) that there is a rod influence in 2° foveal V_λ determinations. While it appears that there is a rod influence in 2° data, even at quite high intensities,⁽³⁵⁾ the result does not follow from the above comparison. LaDriere⁽⁵⁶⁾ found the same type of increase in blue sensitivity relative to red as field size was increased through the values $5'$, $25'$, $50'$, $75'$, and $100'$. Further, this effect was complete when field sizes of $75'$ to $100'$ were reached. (Compare curves 4-5, 11-12, 29-30, 34-35, and 36-37 in Fig. 23.) Thus the effect of increased blue sensitivity with increasing field size terminated within the rod-free area⁽⁹⁾ of the fovea and hence could hardly be attributed to the rods. The results of Ishak,⁽⁵¹⁾ Bedford and Wyszecki⁽⁴⁹⁾ and LaDriere⁽⁵⁶⁾ all indicate that rod effects are absent within $75'$ to $80'$ of the central fovea, in good agreement with the anatomical findings of Polyak.⁽⁹⁾

* See Refs. 52, 62, 64, and 69 on pages 168 and 169.

The relative decrease in blue sensitivity with decreasing field size has been taken as evidence that the very center of the fovea is tritanopic. This view has considerable support in the metameric data by Willmer and Wright⁽⁷⁰⁾ with a 20' field that was definitely dichromatic. As Willmer and Wright themselves noted, however, the results were obtainable only under rigid fixation conditions. The slightest change in fixation caused an immediate breakdown in the matches. This factor probably accounts for the contradictory reports on the tritanopic question, and raises the possibility that the dichromacy is an artifact of rigid fixation without analog in normal situations. Certainly a relative insensitivity to the blue end of the spectrum cannot be required as evidence of central foveal tritanopia in the general population, since Bedford and Wyszecki⁽⁴⁹⁾ did not find such a systematic result with their four subjects.

Furthermore, small-field tritanopia receives a blow from the relation between small-field size and luminance level found by both Bedford and Wyszecki and LaDriere.⁽⁵⁶⁾ The relative decrease in blue sensitivity in going to a smaller field may be reversed by an increase in the intensity level. Finally, the extensive results of Hartridge^{(71),(72)} for small-field stimuli exhibited dichromatic effects throughout the retina, both tritanopic and protanopic.

The preceding results lead to the conclusion that small-field dichromacy is random as to occurrence, type, and location on the retina of a given subject. The possibility of its appearance seems connected with the absence of a sufficient number (or variety) of cone receptors in small fields, and suggests an integrative and/or comparison role at the bipolar level in the color mechanism.

Implications for the Visual Mechanism

It seems reasonable to presume that irregularities in V_λ are manifestations of the underlying cone-receptor mechanisms. It has been believed that these irregularities should provide a means for determining individual receptor response curves. But all such attempts have failed. Hurvich and Jameson^{(64),(69)} tried to fit their V_λ data with a 3-receptor theory without success, and further concluded that V_λ could not be

due to a simple summation of a few chromatic response curves. Rather, their adaptation results indicated that a single mechanism was primarily responsible for the brightness response, with secondary effects superimposed from separate chromatic mechanisms.

The attempts by Boynton⁽⁶³⁾ and Boynton, Kandel, and Onley⁽⁴⁷⁾ to fit experimental V_λ curves with a summation of three of four hypothetical primary sensitivity functions determined by chromatic adaptation results were also unsuccessful. Their best fits with either 3-receptor or 4-receptor assumptions did not describe their data with statistical accuracy. Furthermore, their derived spectral response functions could not be directly related to color-mixture results. Stiles' response functions,⁽⁷³⁾ which were derived from color-mixture data, would not fit the V_λ results of Boynton, Kandel, and Onley.⁽⁴⁷⁾ In fact, Stiles' functions also failed in their intended purpose--namely, a quantitative reproduction of his own experimental results. Further, Stiles stated that for one of his subjects, "it can be shown that complete agreement could not be reached by any change in which the three fundamental response functions remained always positive, linear forms in the CIE distribution coefficients."⁽⁷³⁾

With regard to additivity of luminances (Abney's Law, Eq. (1)) and linear trichromacy (Grassman's Laws, Eq. (14)), Stiles and Burch⁽⁷⁴⁾ found that calculated V_λ curves based on brightness-match data and assuming the validity of Eqs. (1) and (14) differed widely from the measured V_λ functions. Stiles had previously reported⁽⁵³⁾ that the concept of the relative sensitivity function as a summation of three primary processes was not supported by his metameric data. He had concluded a decade before⁽⁷⁵⁾ that the evidence then accumulated on differential adaptation effects ruled out the concept of three independent cone mechanisms with fixed response curves.

Very recent work by Boynton⁽⁷⁷⁾ reinforces this conclusion that 3- or 4-receptor theories exhibiting linearity and summation simply cannot account for the V_λ data. Boynton has followed Stiles⁽⁷⁸⁾ in abandoning the term "receptor" in theoretical analyses of such data, substituting instead the term "fundamental mechanism." This semantic

change is to emphasize that such data will never receive explication in terms of "photopigments" or "cone types" alone. These peripheral "mechanisms" are nonlinear and exert dynamic control on gain and temporal and spatial summation.⁽⁷⁷⁾ The number of such mechanisms postulated to explain the data has oscillated between the Young-Helmholtz magic 3 and a maximum of 7. The most recent attempts by Ikeda⁽⁷⁹⁾ and by Boynton, Ikeda, and Stiles⁽⁸⁰⁾ to experimentally determine these "fundamental chromatic mechanisms" have succeeded in providing data requiring even more complex mechanisms for their explanation.

Thus the vast accumulation of both luminosity and metamerism data in the past twenty years, together with the most skillful attempts to correlate them, have served to reinforce Stiles' original conclusion: Abney's Law, Grassman's Laws, and the Young-Helmholtz concept of three independent cone mechanisms with fixed response curves work fairly well only under certain special conditions, and even then only if the "visual system" is completely adapted to them.⁽⁷⁵⁾ These "laws," which are the heart of the CIE system of colorimetry, cannot be taken as representative of the human visual mechanism in any fundamental sense. Attempts to "explain" the proliferation of results on human color vision on the basis of these concepts are doomed to failure. Yet such attempts are still being made.*

CONCLUSIONS

In summary, a consideration of retinal sensitivity as displayed through relative luminosity and metamerism data seems to establish the following general conclusions:

1. Irregularities in V_λ curves may be an anatomically based characteristic of an individual subject, but their exact origin is not evident from these data alone.

*The above conclusions imply nothing about the existence of human cone photopigments, to be discussed in Section V. They do imply that the recent conclusions by Wald⁽⁷⁶⁾ concerning the "mechanisms" mediating human color vision are not supported by available experimental results.

2. The brightness response mechanism is neither completely separate from the chromatic mechanism nor a simple summation of three or four chromatic response functions.
3. A group of contiguous cone receptors is required for normal hue discrimination, and there is a lower limit in size (area) below which this group cannot function.
4. Linearity in luminance and metamerism is the exception rather than the rule insofar as the human visual mechanism is concerned.

For the purposes of photometry and colorimetry, these results strongly imply that a simple, analytic function would have as much claim to representative validity as the present CIE 1924 Photopic Standard, as has been suggested by Moon and Spencer.⁽⁸¹⁾ In 1955, the CIE recognized^{(25),(32)} the need for revision in both its photopic relative luminous efficiency function and the standard color-mixture functions, and called for new color-mixture data under prescribed conditions,⁽²⁵⁾ including the obviation of Abney's Law. In view of the preceding discussion of luminosity data, it is likely⁽³²⁾ that a new CIE Photopic Luminous Efficiency Standard will be a weighted sum of the new color-mixture functions,⁽⁷⁴⁾ chosen so as to provide the wavelength function best suited for photometric use. According to Judd,⁽³²⁾ "In any case most of the data so far obtained either on color mixture functions or on the luminous efficiency function are likely to be rendered obsolete" so far as photometry and colorimetry are concerned.

V. RETINAL PHYSIOLOGY

The experimentally established spectral responses of the two retinal receptor systems in gross, and the qualified trichromacy of color-matching data, must receive explication in an acceptable theory of human color perception. But they do not, in themselves, serve to determine the mechanisms of cell photoreception and synapsing. These mechanisms have their genesis in the peripheral (retinal) and central (cerebral) portions of the visual system. While proposed mechanisms may be, and have been, advanced without a priori recourse to established physiological facts, it seems unwise and inefficient to do so. In the present section, available experimental data are interpreted for their contributions to some understanding of initial photoreception and of the anatomy, histology, morphology, ontogeny, and electrophysiology of the retinal neurons. Cerebral physiology will be the topic of Section VI.

RETINAL PHOTORECEPTION

The most widespread assumption to date is that the primary mechanism for spectral discrimination consists of several spectrally selective photolabile pigments situated in the outer segments of the receptor cells. If this were the case, then it should be possible to extract such pigments and compare their absorption spectra with pertinent luminosity and color-mixture data. For humans, such a procedure has been successful only for the rods. The search for cone photopigments has given rise to additional methods, the determination of "difference spectra" from in vivo and in situ "bleaching" experiments. Rod photoreception will be discussed first.

Rod Pigments

The most direct connection between luminosity data and a photoreceptor mechanism is that for scotopic vision, mediated by the rods. In essence, it has been shown that the human rod outer segments contain a photopigment, rhodopsin, that is bleached upon exposure to light.

This rhodopsin has been extracted and has an absorption spectrum in good agreement with the human scotopic luminous efficiency function. While this was shown originally by König in 1894 and verified by Crescitelli and Dartnall⁽⁸²⁾ in 1953, completely conclusive evidence was not furnished until 1958 by Wald and Brown.⁽⁸³⁾

The results for the rod mechanism are not of primary concern in this study, since they could be applicable to the cones and color vision only by analogy. But a brief summary is necessary to put the search for cone photopigments in proper perspective. Those interested in further information should consult Crescitelli⁽⁸⁴⁾ on rod photopigments, and the papers by Wald,^{(34),(85)} Moon and Spencer,⁽⁸⁶⁾ Kropf and Hubbard,⁽⁸⁷⁾ and Weale⁽⁸⁸⁾ on the chemistry of rod vision.

Figure 28 shows a comparison of the spectral sensitivity of human scotopic vision with the absorption spectrum of human rhodopsin as determined by Wald and Brown.⁽⁸³⁾ The absorption spectrum is for the rhodopsin pigment in a suspension of human rod particles,⁽⁸³⁾ rather than a rhodopsin extract. Measurements on the latter^{(82),(83)} result in a slight shift of the absorption maximum to a shorter wavelength⁽⁸³⁾ due to the extraction technique. The scotopic luminosity data for the aphakic human eye in Fig. 28 are those of Wald,⁽³⁶⁾ adjusted to an equal quantum intensity basis⁽⁸⁹⁾ for direct comparison with the absorption results. The solid curve is a similarly quantized average scotopic luminosity curve for normal subjects (see Fig. 19), corrected for transmission of the ocular media.⁽⁹⁰⁾ The aphakic data represent the closest one can approach to the intrinsic sensitivity of the rods in vivo, since the primary colored structure of the eye, namely, the yellow lens, has been removed in the cataract operation.

The results in Fig. 28 seem conclusive: The photopigment rhodopsin mediates scotopic (rod) vision by virtue of the photochemical changes it undergoes upon exposure to radiation. This conclusion is reinforced by the fact that estimates of the rhodopsin concentration in situ based on extraction,^{(82),(83)} reflection densitometry in the living eye,^{(88),(91)} transmission densitometry for small areas in retinal preparations,⁽⁹²⁾ and transmission densitometry in single rod elements⁽⁹³⁾ are all in reasonable agreement. It is important to remember

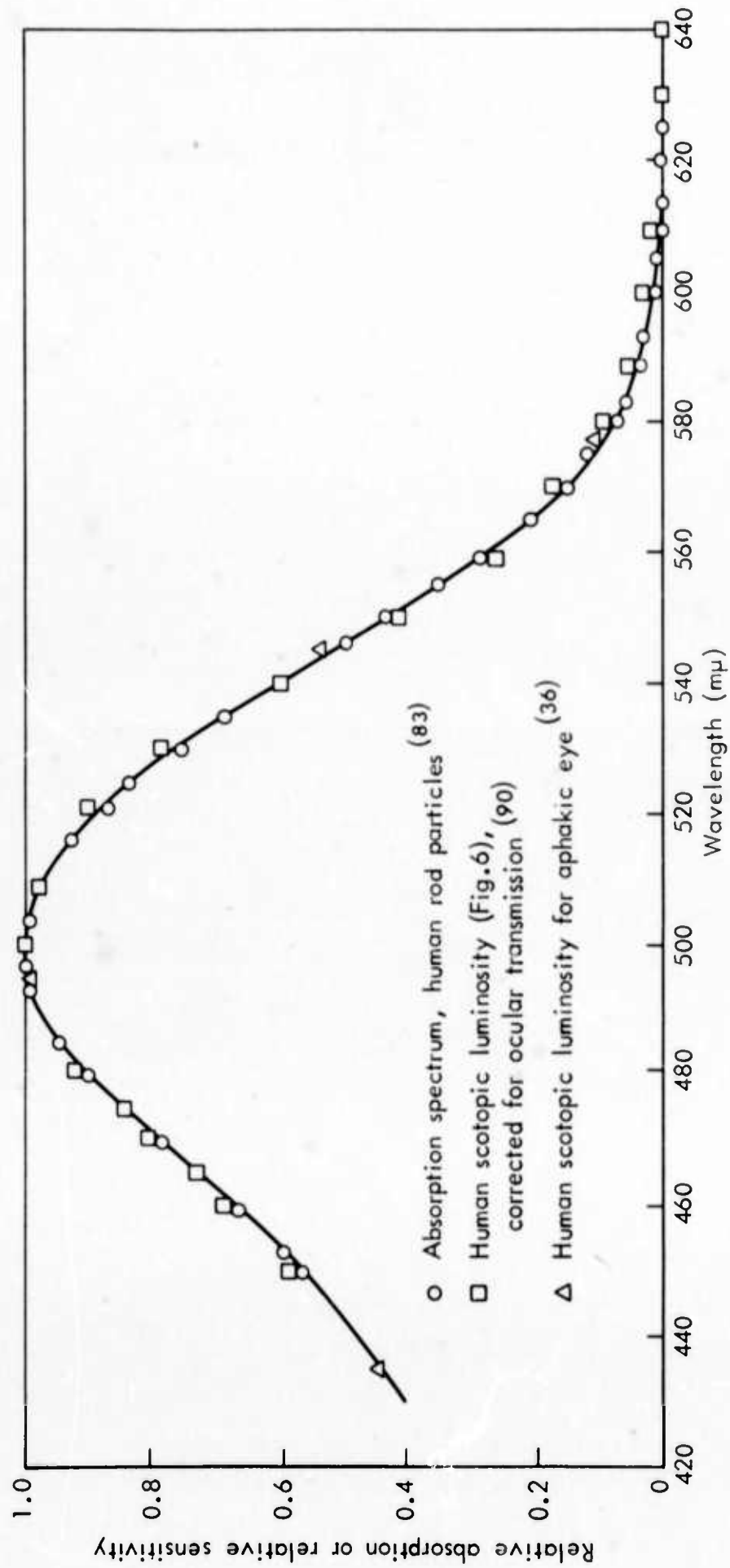


Fig. 28—Comparison of human rhodopsin absorption spectrum and the spectral sensitivity of scotopic vision

this basic agreement in rod pigment results among the several methods when interpreting the data from the same methods applied to the cones.

Cone Pigments

The qualified trichromacy of metameric data could be rather simply explained by the existence of three suitable photopigments distributed among the cones, in analogy with the rod mechanism. However, extending such an explanation to human color vision in general would be too simple, as noted in Section IV. Nevertheless, the search for evidence to support this basic Young-Helmholtz concept still proceeds, and we turn now to a consideration of the results to date. In essence it may be said that the preponderance of experimental data still fails to give unequivocal support for such a hypothesis. The experimental procedures used in the search are extraction and in situ and in vivo spectrophotometric densitometry, as with the rods.

The only visual photopigment that has ever been conclusively recovered from primate retinas by extraction is rhodopsin, presumably entirely from the rods. No specifically cone pigment has been found by this method. In fact, the only visual pigment extracted from any vertebrate retina with a claim to be regarded as a cone pigment is iodopsin, extracted by Wald⁽⁹⁴⁾ from the chicken retina. Such extracts contain a mixture of two photopigments: rhodopsin, with an absorption maximum near 500 mμ, and iodopsin,⁽⁹⁴⁾ with an absorption maximum near 560 mμ. Since the chicken retina has a mixture of both rods and cones inseparable in the extract, if one assumes that the rhodopsin is an exclusive resident of the rods, then iodopsin is thereby a cone pigment. This inference has additional support for the chicken as well as for certain other birds and for turtles. Their cones contain colored light-stable filters between the inner and outer segments. These filters, together with the single photopigment iodopsin, would provide a possible mechanism for hue discrimination.⁽³⁴⁾

The amount of iodopsin found in the chicken extract was small. Despite the great preponderance of cones in the chicken retina, the extracts contained larger amounts of rhodopsin than of iodopsin.⁽³⁴⁾

Wald concludes from his studies that the concentration of iodopsin in a single cone is two to three orders of magnitude less than that of rhodopsin in a rod.⁽³⁴⁾ Several attempts by Wald to extract photopigments from the all-cone turtle retina were unsuccessful.⁽³⁴⁾ Dartnall⁽⁹⁵⁾ has reported the successful extraction of a pigment from the pure cone retina of the grey squirrel. The pigment appears to be a rhodopsin type, i.e., based on retinene₁, with an absorption maximum near 500 mμ. But the extract concentration was too low for a completely meaningful analysis. Moreover, comparison of in vivo bleaching measurements on the squirrel retina by Weale⁽⁹⁶⁾ with the amount of pigment extracted by Dartnall resulted in gross disagreement. The inferred in situ density of the pigment from the bleaching measurements was two orders of magnitude higher than that from the extraction results.⁽⁹⁵⁾ One should recall that such comparisons with the rod pigment, rhodopsin, are in good agreement. Further, there are no apparent reasons to presume that the chicken iodopsin found by Wald, or the rhodopsin-like squirrel pigment found by Dartnall, is a cone photopigment of the human eye. Finally, it is important to note that the previously discussed extracts of human retinæ by Crescitelli and Dartnall⁽⁸²⁾ and by Wald and Brown⁽⁸³⁾ did not show the presence of any photopigment other than rhodopsin.

Rushton,⁽⁹⁷⁾ using his double pass absorption technique in vivo, felt that he had located at least two cone pigments in the living human eye. But his technique and interpretation⁽⁹¹⁾ of data, and explanation⁽⁹⁸⁾ of his result that the in situ cone photosensitivity is several times that of rhodopsin despite the fact that no human cone pigment has ever been found by extraction, leave his conclusions open to serious challenge.⁽⁹⁹⁾ Rushton, himself, has recently repudiated his earlier conclusions with a vivid illustration of the extreme difficulties involved in interpreting retinal reflection densitometry.⁽¹⁰⁰⁾ In a recent series of papers,⁽¹⁰¹⁻¹⁰³⁾ Rushton has given a detailed analysis of his technique and its application to a protanope. His arguments that the in vivo densitometric results imply the existence of two cone photopigments, despite continued discrepancies in pigment concentration results, are unconvincing. Similar measurements have been reported and analyzed by Weale.⁽⁸⁸⁾ Both Rushton and Weale postulate

histological details of cone structure to provide geometrical enhancement of the effectiveness of a minute amount of photopigment and thereby "explain" the discrepancy between their results and those by extraction. It appears much more reasonable that these in vivo densitometric results are a biophysical analog of the psychophysical luminosity and metameric data. Neither one nor both constitute sufficient evidence for human cone photopigment specification. Neither precludes alternative explanations.

Cone Spectral Sensitivities

Very recently, microspectrophotometric techniques have been developed that permit transmission densitometry measurements on single human cone receptors in retinal preparations. The initial results for several human and simian cones have been reported by Marks, Dobelle, and MacNichol⁽¹⁰⁴⁾ and for four human cones by Brown and Wald.⁽⁹³⁾ The data consist of absorption difference spectra obtained by subtracting a densitometer trace taken after bleaching the receptor from a trace taken before the bleaching exposure. This procedure results in typical broadband curves similar in shape to that of rhodopsin in Fig. 28. The peak in the absorption difference spectrum ranges from 435 to 585 m μ for the curves reported in the two investigations above.^{(93),(104)} Both groups of authors proposed that their results indicate three types of cone receptors: a blue, peaked between 440 and 450 m μ ; a green, peaked between 520 and 540 m μ ; and a "red," peaked between 550 and 580 m μ . However, this suggestion is obviously not substantiated by the limited number of primate cones examined so far.

The basic idea behind the microspectrophotometry of single receptor cells is to determine by population studies whether there are three spectrally distinct photopigments in the cones, or three broad classes where members of each class are similar but not identical in absorption spectrum. Thus the difficulties of extraction could hopefully be sidestepped temporarily. (The assumption of "three" pigments or "three" classes, above, is based on the working hypothesis that the approximate trichromacy of metameric data is best explained by such an assumption.)

A histogram of 113 single cones from the goldfish retina convinced Marks⁽¹⁰⁵⁾ that the goldfish has three distinct cone photopigments with absorption maxima at $455 \pm 15 \text{ m}\mu$, $530 \pm 5 \text{ m}\mu$ and $625 \pm 5 \text{ m}\mu$. The blue pigment had a secondary peak in the red. Cone pigment concentration in the goldfish outer segments was found to be about the same as that of rhodopsin in frog rods. These conclusions are shared by Liebman and Entine⁽¹⁰⁶⁾ on the basis of similar measurements on 60 goldfish cones. The desirability of applying such a successful technique to the primate retina is obvious. However, the limited results obtained so far do not follow the pattern exhibited by the goldfish data.

The immediate question then is, What do the above-reported primate difference spectra actually represent? The authors assume that the curves, after "appropriate" corrections for any bleaching, are indicative of several different cone photopigments. Such an assumption involves the negation of any alternative explanation, e.g., that the spectral selectivity of an individual cone might be due primarily to its physical characteristics (size, index of refraction, etc.) rather than to any photopigment that it might contain.

A possible alternative explanation might lie in a wave-optics analysis of the spectral energy density in cone outer segments as a function of their physical characteristics. The difficulty here is that to be directly related to the above receptor "sensitivity" measurements, such an analysis would be exceedingly complex. To obtain tractable boundary conditions, one would have to assume an extremely simple model for the cone receptor.

Snitzer and Osterberg⁽¹⁰⁷⁾ have investigated multimodal waveguide propagation in clad dielectric cylinders with refractive index differences comparable to those in human retinal structures. Figure 29 shows some of the direct image patterns that they obtained with cylinder diameters comparable to human cone receptor geometry. The existence of such waveguide effects in human cone outer segments has been demonstrated by Enoch⁽¹⁰⁸⁾ and is discussed later in this section under "Receptor Histology" (see page 83).





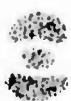

Image						
Mode*	TE ₀₁	TM ₀₁	HE ₁₁	HE ₁₂	HE ₁₂ +EH ₁₁	2HE ₁₂ +EH ₁₁ +HE ₃₁
Radius (μ)	0.83	1.32	0.53	1.32	1.32	1.32
Wavelength (μ)	0.603	0.491	0.548	0.548	0.550	0.548

Fig.29—Sketches of direct images of waveguide modes and mode mixtures in optical fibers (Ref. 107)

For present purposes it is sufficient to note that the undeniable existence of waveguide effects in individual cone receptors implies that the "difference spectra" measured by Marks, Dobelle, and MacNichol⁽¹⁰⁴⁾ and by Brown and Wald⁽⁹³⁾ are not simply related to cone photopigments. If one denotes by F the fraction of incident energy available to an absorber inside the cone, then their "dark spectra" are proportional to $F(1 - \alpha_1)$, where α_1 is the total absorption on passage through the cone. Similarly, their "bleached spectra" are proportional to $F(1 - \alpha_2)$, if one assumes that the photochemical changes resulting from energy absorption do not change the waveguide characteristics. Under these conditions, the "difference spectra" that they reported are actually proportional to $F(\alpha_2 - \alpha_1)$. Now F is a complicated function of wavelength, location in the receptor, receptor geometry and physical characteristics, modes propagated, and methods of excitation. Yet the F 's must be determined if the data are to yield simple "difference spectra" as intended. Finally, it is recalled that even the "difference spectra"

*For terminology, see Ref. 119.

themselves are not direct determinations of primary cone pigment absorptions. The bleaching products must also be determined.

The need for much more extensive primate results is obvious. In the meantime, the caution required in the interpretation of presently available data will be illustrated with the following comments on the Marks, Dobelle, and MacNichol⁽¹⁰⁴⁾ results:

1. Available spectra for primates do not constitute sufficient data for a meaningful histogram.
2. The question of waveguide effects in primate cones continues unanswered. Goldfish cone outer segments are broad, stubby, and "cone" shaped. The microspectrophotometric measurements were performed with the test beam normal to the goldfish cone axis. Geometrical considerations would indicate that waveguide effects were not an important factor in these measurements. On the other hand, primate cone outer segments are long, slender cylinders, and the measurements on them were made with paraxial test beams. Waveguide effects may be extremely important here.
3. All the primate measurements so far have been on peripheral cones, from regions roughly 15° to 30° from the fovea.⁽¹⁰⁹⁾ This decision was dictated by the capability of experimental techniques. Foveal and parafoveal primate cones are even more slender and much more elongated, enhancing any waveguide effects. Of course, color discrimination is highest in the fovea.
4. The data for the primate cones so far examined definitely do not follow the pattern exhibited by the goldfish results. The three "receptor types" found in fish have spectral sensitivity curves that are well spaced across the visible spectrum, with a "Young-Helmholtz" type red receptor peaked at 625 m μ . The primate cone spectra are not so simply separated, and a true "red" receptor was not reported.

In any event, the existence of waveguide effects in human cone outer segments, coupled with the obviously present geometrical variations, constitutes a major handicap to proponents of a 3- (or 4-)

"receptor" theory. All quantitative applications of this concept involve three (or four) fixed, broadband, overlapping, spectral response curves, with the resulting hue discrimination dependent on very small differences in the responses of the 3 (or 4) receptors. Even with rigidly fixed pigment responses, a waveguide effect that varied with cone physical characteristics could completely swamp such a discrimination mechanism (although metamerism could still be explained thereby).

Thus the question of the number and role of cone photopigments in the human visual system must await further research for a definitive answer. In the meantime, the implications of the goldfish results discussed previously should not be allowed to restrict the primate mechanisms considered as feasible. One must remember Gordon Walls' conclusion of twenty years ago, recently echoed by DeValois,⁽¹¹⁰⁾ that there is good evidence that color vision systems have evolved several separate times--once in insects; again in fish, reptiles, and birds; and still again in primates. It would thus be little more than a coincidence if the details of the visual systems in man and fish were the same.

The principal points in this discussion of receptor sensitivity and visual pigments may be summarized as follows:

1. The role of rhodopsin in the rods as mediator of human scotopic vision is established. The measurements on human rhodopsin by extraction, in situ methods, and in vivo methods are in mutual agreement.
2. No primate visual pigment other than rhodopsin has been found by extraction. The evidence by extraction for a single cone pigment, iodopsin, in certain lower vertebrates is tenuous. No evidence by extraction exists for a multiplicity of cone photopigments.
3. The estimates of cone "pigment" concentrations based on in situ and in vivo measurements are not in agreement. Both methods indicate concentrations that should be detectable by extraction.
4. The limited number of spectral sensitivity measurements on individual human cone receptors tends to indicate a multiplicity

of "types." Such measurements do not necessarily imply the existence of an analogous variety of cone photopigments. The spectrally selective waveguide effects on such data must first be determined.

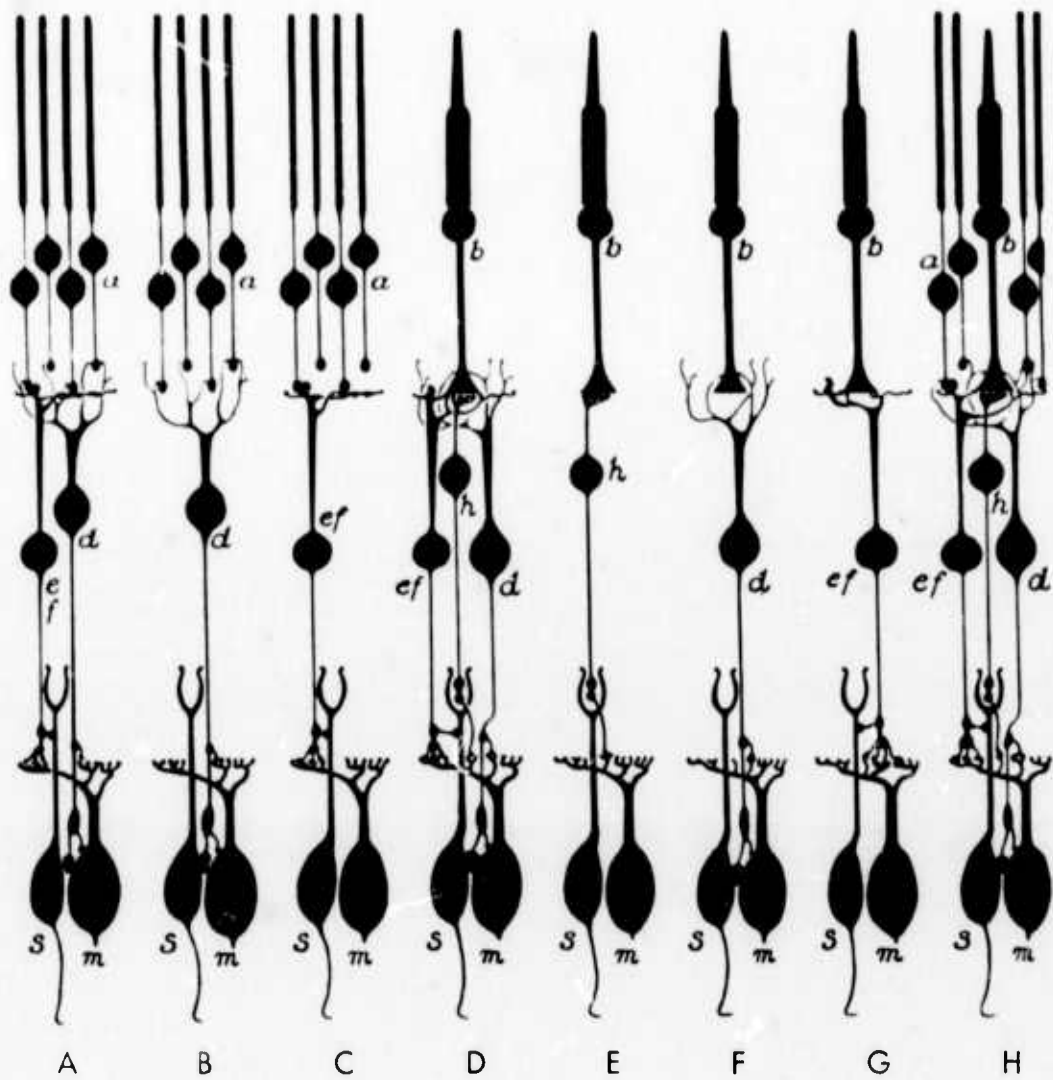
5. The principal evidence supporting the widespread belief in three (and only three) human cone photopigments is still not extraction or densitometric measurements but the approximate trichromacy of metamerism, as previously noted by Boynton.⁽¹¹¹⁾ Such a belief still does not necessarily imply that the "fundamental sensitivity curves" of human color vision are those of photopigments. The conclusions of Section IV and the present discussion indicate that even if one had a complete specification of three cone pigments, most of the problems about human color discrimination would remain unsolved.

RETINAL ANATOMY

In view of the preceding conclusions, the anatomical details of retinal structures receive enhanced importance both in the initial photoreception and in the peripheral processing of the receptor excitations.

One turns first to Polyak,⁽⁹⁾ the most quoted reference in retinal anatomy, appealed to by almost all theorists in color vision. It is therefore astonishing to note that not only has the present writer found but a single specific reference to what is probably the most germane result in Polyak's voluminous work as regards color vision, but he has recurrently found Polyak cited as proof of the exact opposite. The exception noted, a paper reporting data substantiating Polyak's findings, is discussed below. The result is concerned with the types and synaptical relationships of the retinal neurons that constitute the visual pathway--the rod and cone cells, the bipolar cells, and the ganglion cells.

One of the most widely reprinted figures in Polyak's book is shown here as Fig. 30. It represents the grouping of the neurons of the primate retina into functional systems; and it must be emphasized that Fig. 30 represents just that: types of synaptical contacts, not receptor cell anatomy.



- A = rod system
 B = reduced rod system with mop bipolar
 C = reduced rod system with brush or flat top bipolar
 D = complete cone system
 E = reduced cone system with midget bipolar
 F = reduced cone system with mop bipolar
 G = reduced cone system with brush or flat top bipolar
 H = mixed complete rod and cone system

- a = rod
 b = cone
 d = mop bipolar
 e = brush bipolar
 f = flat top bipolar
 h = midget bipolar
 m = any diffuse ganglion
 s = midget ganglion

Fig.30—Grouping of neurons of the primate retina into functional systems by Polyak (adapted from Ref. 9)

The legend of Fig. 30 should be read for essential details and terminology. The neuron grouping of greatest interest is D, the complete cone system with no rods present. Each cone in the human retina is synaptically related to three (h, e, f) bipolars and possibly also the fourth (d).⁽⁹⁾ This interconnection alone gives at least three (b-h, b-e, b-f) and possibly four (b-d) "distinct channels into which the excitation from a single cone may proceed."^{*} If each bipolar variety is capable of responding dynamically only in its own specific way, a single cone excitation may in this way elicit three or four distinct dynamic reactions on the bipolar level.

When the connections of the same bipolars to the ganglion cells are examined, it is found that each midget ganglion (s) is related to every bipolar variety, "duplicating their cone relationships, i.e., (h-s), (e-s), (f-s) and (d-s) synapses."^{**} Here, then, the effect is reversed. The three or four distinct bipolar channels converge onto a single midget ganglion cell channel whose axon runs through the optic nerve to the brain proper. Simultaneously, all four bipolar varieties synapse with a diffuse ganglion (m, Fig. 30).

In Polyak's words again, "these are the actual anatomical facts." But they are precisely opposite the widely held view that the cone-bipolar-ganglion synaptical relationship is always afferently convergent. It has been widely assumed that in the afferent pathway the number of cones is larger than the number of bipolars, which is in turn larger than the number of ganglions. Others have assumed that in the center of the fovea there is a one-to-one relationship among the three types of neurons, i.e., only a (b-h-s) pathway. Both assumptions have been supported by a reference to Polyak, yet he rather clearly contradicts both views: "In most of the retina, and particularly so in the central area, each cone is related to a single, individual h-bipolar and at the same time to a group composed of the two or three diffuse varieties of bipolar cells (d, e, f, g)."^{***}

* See Ref. 9, p. 572, emphasis supplied.

** Ibid., emphasis supplied.

*** See Ref. 8, p. 249, emphasis supplied.

Polyak's statements have explicit experimental verification in the striking histological measurements on the human retina by Vilter.⁽¹¹²⁾ Vilter examined microscopically a mounted 2- μ -thick cut running from the dorsal to the ventral aspect of a human retina, and counted the number of retinal cells (rods, cones, bipolars, and ganglions) appearing in successive linear fields of 100 μ each. The count was continuous from the central fovea to the termination of the receptor area of the retina, the ora serrata, which corresponds to a total visual field of about 80°. His results are shown in Fig. 31. The left ordinate scale gives the number of neurons in a linear field of 100 μ . The abscissa represents successive fields of 100- μ extent running from the retinal periphery to the central fovea. The cone cell (c) counts are plotted individually. Each of the other neuron distribution curves is similarly determined by 140 points.

To aid in the interpretation of the cone-bipolar-ganglion synaptical relationships in the central fovea, where the cone synapses are laterally displaced, Vilter also counted the number of cone synapses when these began to outnumber the cones as the fovea was approached. This count is shown as a separate curve in Fig. 31.

From Fig. 31, one notes that in the outer periphery, the cone and ganglion distributions are practically constant. There are about three cones and one ganglion in each 100- μ field from 80° up to about 35° from the fovea centralis. More importantly, in this region the increase in rods per unit retinal area is followed closely by a corresponding increase in the number of bipolars. This correlation is quite precise with a rod/bipolar ratio near 5:1. If each cone is assumed to retain a midget bipolar (h, Fig. 30) after Polyak,⁽⁹⁾ then the remaining bipolars would presumably be of the mop variety (d) always found with the rods.⁽⁹⁾ In this case, the ratio of rods to rod bipolars is highly constant near 6:1 over the region from 70° to about 35°.

From about 35° on toward the fovea centralis there is a distinct change in the neuronal ratios. The number of cones and ganglions per unit retinal area begin to increase from their constant peripheral values. The number of rods shortly reaches a peak, then decreases rapidly. The previously simple correlation in neuronal ratios no longer holds.

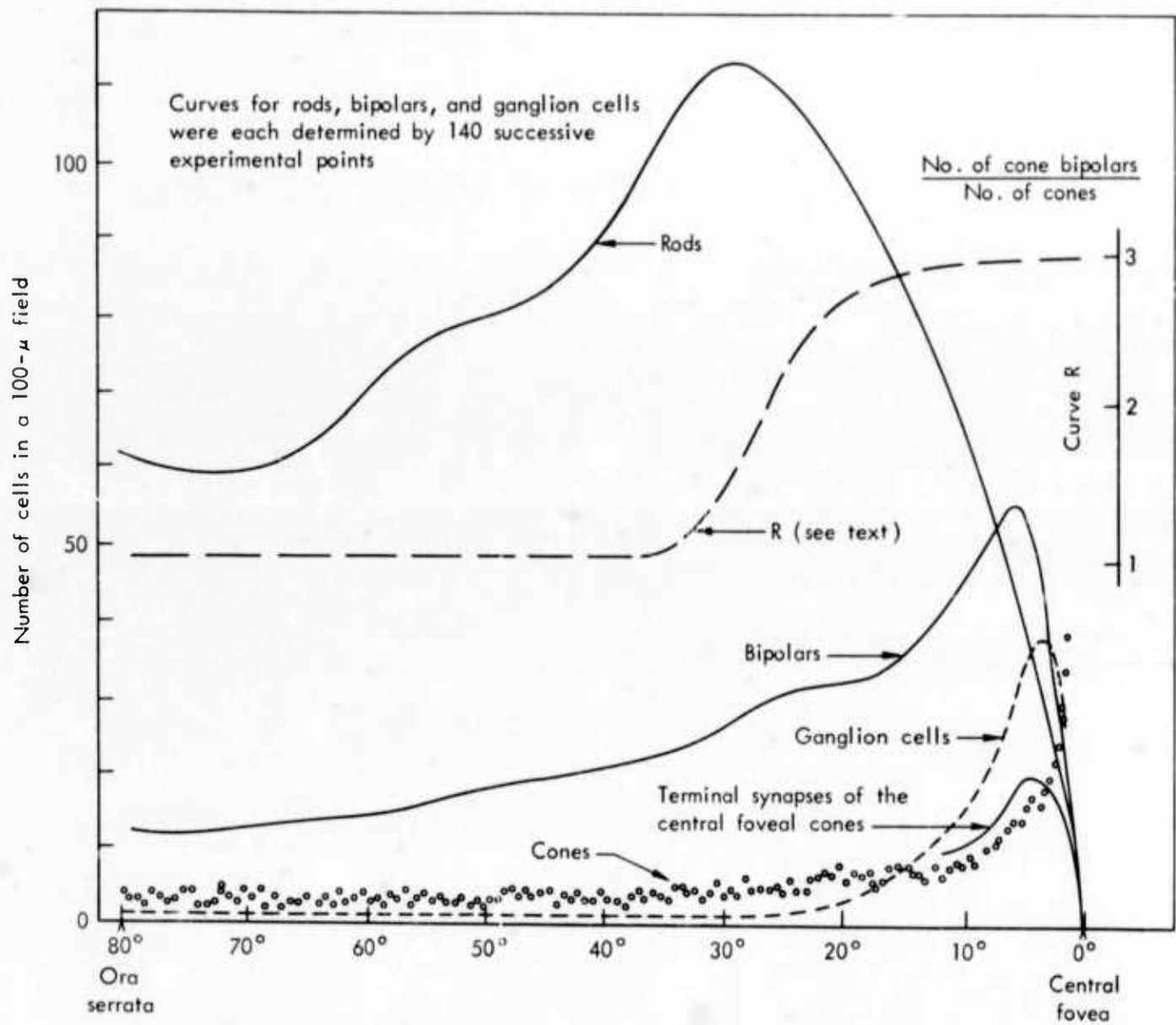


Fig.31 — Distribution of neurons in a 2- μ -thick dorsoventral cut through the human retina, extending from the central fovea to the ora serrata by Vilter (adapted from Ref. 112)

But if the rod to rod-bipolar ratio established over a range of 45° of the periphery is assumed to continue to hold, the remaining "cone bipolars" are monotonically related to the cones. This ratio of cone bipolars to cones is shown as curve R in Fig. 31. The ratio increases from 1:1 at 35° , through 2:1 at 28° , to a plateau of 3:1 at 15° from the fovea centralis. Figure 31 also shows that in this central region the ganglion/cone-pedicle ratio increases to a plateau value of 2:1.

The implications of the joint Polyak-Vilter results are clear. In the central retina, where color vision is fully developed, each cone has three bipolars directly associated with it. One of these is the midget (h) bipolar monosynaptically related to its cone. The other two are the brush (e) and flat-top (f) bipolars, which synapse both with their parent cone and with the receptor cells immediately adjacent. This would result in a reciprocal overlap of the receptor fields covered by the brush and flat-top bipolars, precisely as stated by Polyak.*

It must be emphasized that the above reference to the mop (d) bipolar as a "rod bipolar" does not imply that it is synaptically related only to the rods. On the contrary, the mop bipolar dendrites "indiscriminately envelop both kinds of photoreceptors."** What is meant is that the mop bipolars are absent in the cone synapse region corresponding to the rod-free central fovea,*** and hence their role in hue discrimination is problematical. For the central fovea, only the cone (e, f, h) varieties of bipolars are found.

The correlation of the results in Fig. 31 with a typical psychophysical mapping of the "color zones" in the human retina is striking. Figure 32 shows such a mapping. Selecting either of the radii corresponding to the ora serrata at 80° for comparison with Fig. 31, one notes in Fig. 32 that the retina is achromatic beyond 40° . This corresponds in Fig. 32 to the region where the bipolar/cone ratio is 1:1. Full color vision is reached in Fig. 32 only inside 20° of the fovea centralis, corresponding in Fig. 31 to a bipolar/cone ratio of 3:1.

* Ibid.

** Ibid., p. 268.

*** Ibid., pp. 268 and 279-281.

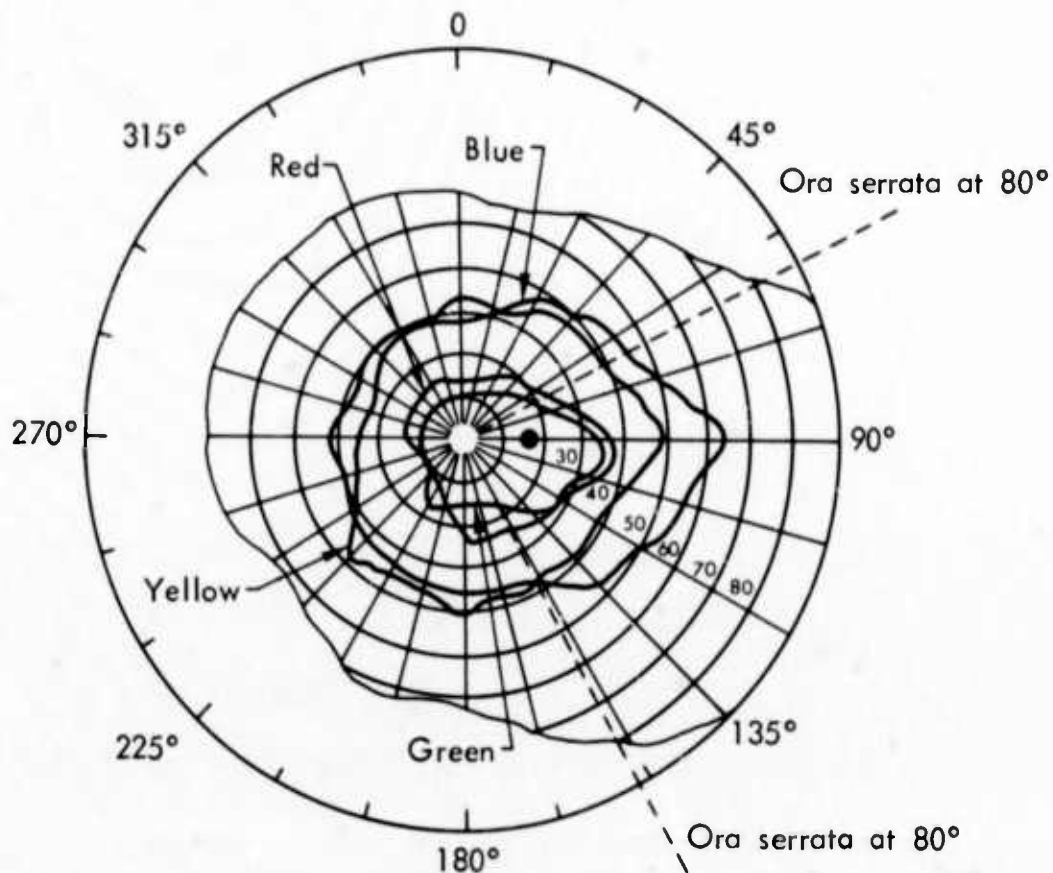


Fig.32 —Correlation between the ratio cone bipolars/cone and the limits of color sensations in the visual field (Ref.1)

Figure 32 further implies that a bipolar/cone ratio of 2:1 suffices only to distinguish the long and short wavelengths of the spectrum.

It is obvious that this "bipolar tripartition" must be considered in any theory of human color vision. Nature simply would not supply us with such anatomical riches without purpose. Further elaboration of the many possibilities is not appropriate here, but it is interesting to note that even the simplest synaptical system is already quite complex. This is illustrated in Fig. 33.

Neglecting all adjacent interactions, a single excited cone may elicit one or several distinct dynamic states in its three directly

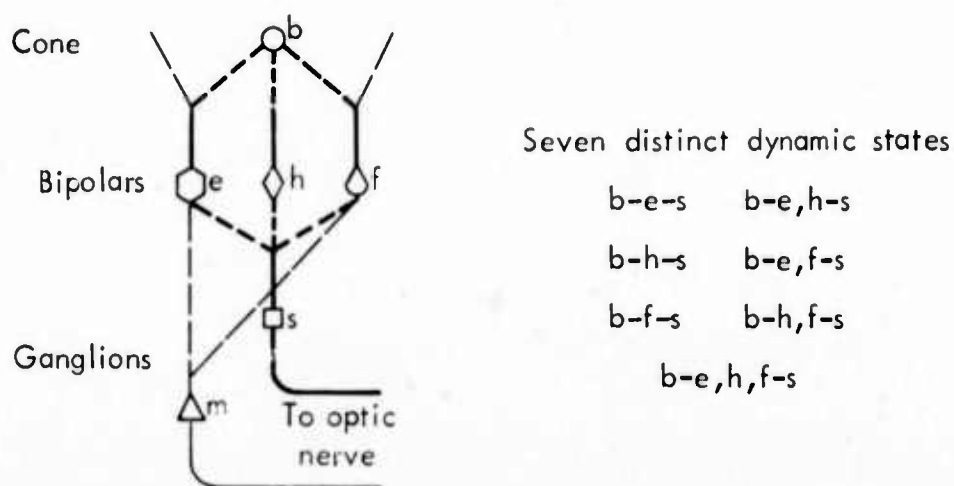


Fig.33—Partial synaptical system with bipolar tripartition

related bipolars. These, in turn, elicit an excitatory reaction in their common midget ganglion (s), which is the final pathway to the brain proper. As the figure shows, there are seven possible distinct dynamical responses in the (s) optic nerve to the excitation of a single cone receptor.

RECEPTOR HISTOLOGY

In view of the aforementioned possibility of waveguide effects on receptor sensitivity, the physical details of individual receptor cells become important. Let us now consider some of the available histological data concerning such cells. Unfortunately, the results for human material are meager.

Index of Refraction

No data for human receptor cells were found in the literature. Barer⁽¹¹³⁾ has summarized some measurements by R. L. Sidman of refractive indices of different parts of rods and cones removed from fresh

retinas of several species (frog, salamander, turtle, chicken, pigeon, rat, mouse, and monkey). The results are shown in Fig. 34. Note that

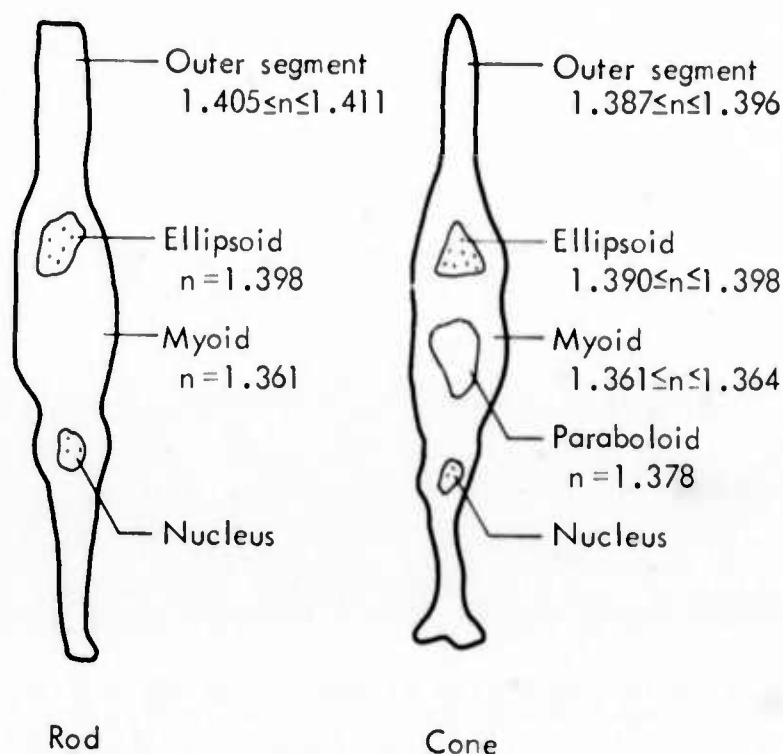


Fig.34—Summary of refractive indices of rod and cone elements
by R. L. Sidman (Ref. 113)

there is remarkably little variation in index of refraction of comparable structures in the different species despite enormous differences in dimensions and mass. As Barer stated, this "suggests that these values conform to some fundamental pattern and are not merely haphazard."⁽¹¹³⁾ Barer also notes that the lower value of n for the cone outer segment (i.e., 1.387) is probably the more correct value.

The determination of the index of refraction of cone outer segments assumes its full importance only with a knowledge of the refractive index of the surrounding medium. Very little is known about this tissue fluid. Barer⁽¹¹³⁾ notes that the lowest value it could have is about 1.334, with a more likely value being about 1.347. The relative refractive index between the cone outer segment and its surround would

thus be about 1.030. But this value would be expected to vary for any given cone outer segment about a mean characteristic of the total cone population for the individual. Barer states⁽¹¹³⁾ that his results for refractive indices of red blood cells from an individual often approximate a Gaussian distribution. Such a variation in refractive index among cone outer segments, coupled with the equally probable variations about means in outer segment geometry, provides a basic mechanism for differences in spectral sensitivity among the cone population without the need for several selective photopigments.⁽¹¹³⁾

Cell Geometry

The geometry of the cone cells thus assumes specific importance with respect to hue discrimination as well as to visual acuity, the latter having long been recognized. Polyak's data⁽⁹⁾ on human rod and cone diameters may be summarized as follows. In the rodless central fovea, the cone inner segment diameters are about 1.0 to 1.5 μ , and outer segment diameters are about 1 μ . The separating neuroglial sheath about the outer segments is about 0.3 μ . Thus the central center-to-center cone distance is about 2 μ or slightly greater.⁽⁹⁾

In the foveal periphery the cone inner segments increase to 3.5-4 μ , but their outer segments are still small. Polyak says the outer segments are about 1.3 μ . The rods here are about 1 to 1.5 μ in diameter. In the periphery near the fovea (the perifovea), only one rod separates two adjacent cones. The rods are typically 1 to 2 μ in diameter, and the cone inner segments have a diameter of 4 to 5 μ . Polyak gives no outer segment values here or further into the periphery.

In the central periphery there are two rods between adjacent cones, the rods being 1.5 to 2.5 μ in diameter and the cone inner segments having diameters of 5 to 6 μ . Further out into the periphery the number of intervening rods increases to three but not to more than four, whence the ratio remains constant to the ora serrata. The rods increase to about 3 μ in diameter, and the cone inner segments increase to 7 or 8 μ in diameter, and possibly to 9 μ at the ora serrata.

All the above data are from Polyak.⁽⁹⁾ They have been questioned frequently by those who feel that the above cone outer segment diameters are too small. O'Brien⁽¹¹⁴⁾ has indirectly confirmed Polyak's values with measurements on human material in the center of the fovea. O'Brien measured the average center-to-center distance for small groups of four to five cones with sections through the cone inner segment region. He found a value from 2.02 to 2.32 μ ($\pm 0.10 \mu$ on each) for the most central cones. This puts an upper limit on the cone outer segment diameters of about 1.4 to 1.7 μ in this region, depending on the thickness of the neuroglial sheath and indicates that Polyak's values⁽⁹⁾ are not too small.*

O'Brien⁽¹¹⁴⁾ gives a very plausible reason why there has been so much uncertainty about cone outer segment diameters in both central and perifoveal areas. The measurements can be made more meaningfully with transverse rather than parallel eye sections. But this renders the material useless for other types of work, and the rarity of available human material has made such sectioning rarer still. O'Brien states that his best measurements were with flat mount, unstained, whole retinas.

The above discussion indicates that data exist for cone inner segment diameter and its variation with retinal location. Also, data exist for cone outer segment diameter in the fovea. But there is still some question about the variation of cone outer segment diameter with retinal location. Unfortunately, there seems to be some confusion in the literature as a result of authors failing to clearly state whether their quoted "cone diameters" refer to inner segment or outer segment. The difference is both considerable and extremely important.

*The very recent electron microscope photographs by Dowling⁽¹¹⁵⁾ of simian cone cells provide the first indisputable data that primate cone outer segments have a diameter of the order of μ or less. Dowling⁽¹¹⁶⁾ has observed individual foveal cones with outer segment diameters ranging from 0.75 μ to slightly over 1 μ , the 0.9- μ -diameter outer segment of the cone published in Ref. 115 being typical on the basis of an examination of less than fifteen cones. Dowling has not yet measured a sufficient number of cones to allow an estimate of the population density versus outer segment diameter. The obvious importance of waveguide effects in such small-diameter "cylinders" makes a population study crucial.

Figure 35 shows Polyak's sketch⁽⁹⁾ of the cone receptors in the very center of the fovea. The most striking feature of this sketch is not the rapid increase in cone inner segment diameter from the center out, but the highly uniform outer segment diameters over the region. This is even more graphically illustrated in Fig. 36, which shows cones "teased" from a region outside the central fovea of a monkey retina. Despite the thick and highly tapered inner segment, the cones have very thin, only slightly tapered, outer segments. These features are also illustrated by O'Brien,⁽¹¹⁴⁾ who presented a sketch of a typical human extra-foveal cone whose outer segment is quite cylindrical with a diameter somewhat less than 1 μ . The cone had an inner segment some 4 to 5 times this value.

These data can be considered as strong evidence that the cone outer segment diameters are highly uniform throughout the entire central area of the retina, exhibiting only a small variation about a mean diameter of the order of 1 μ . This conclusion strongly conflicts with LeGrand⁽²⁾ (as well as with many others), who seems to have misinterpreted Polyak's statements. Enoch,⁽¹¹⁷⁾ quoting the same source as LeGrand, gives values of the cone outer segment diameter as 1.0 μ in the central fovea and 1.3 μ everywhere else. Although much closer to the present interpretation, this is still not a justifiable statement on the basis of Polyak^{(8),(9)} alone. The present writer feels that the conclusion reached herein is reasonably based on experimental evidence and is extremely important to hue discrimination. It seems highly probable that only in the outer periphery, where color vision begins to break down (in a selective and irregular manner,⁽¹¹⁸⁾ it should be recalled), do the cone outer segment diameters depart widely from their mean central value.

Waveguide Modes

The previously mentioned results of Enoch⁽¹⁰⁸⁾ demonstrating that human (and simian) retinal receptors exhibit waveguide modal patterns and other phenomena characteristic of dielectric waveguides⁽¹¹⁹⁾ give very important evidence concerning receptor outer segment histology.

- A few rods are shown near the outer edges; the balance are cones
- Note the uniformity in diameter of the cone outer segments despite the wide variation in inner segment diameter

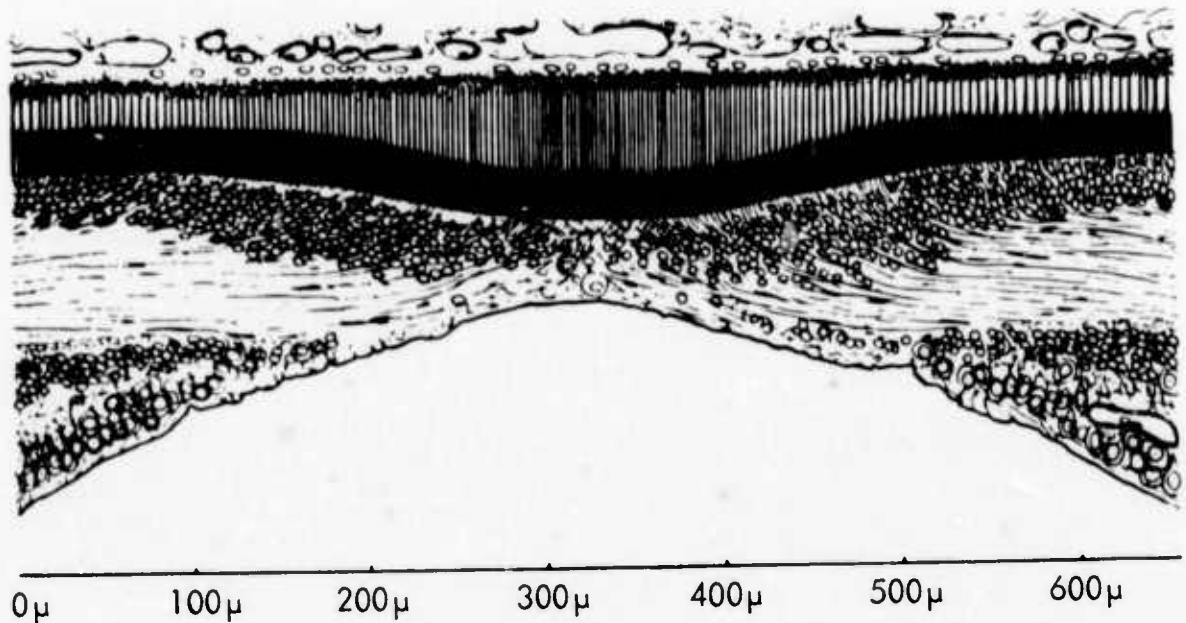


Fig.35—Sketch of the deepest portion of the foveal pit of the human retina by Polyak (by permission from S. Polyak, *The Vertebrate Visual System*, The University of Chicago Press, Chicago, Ill., 1957)

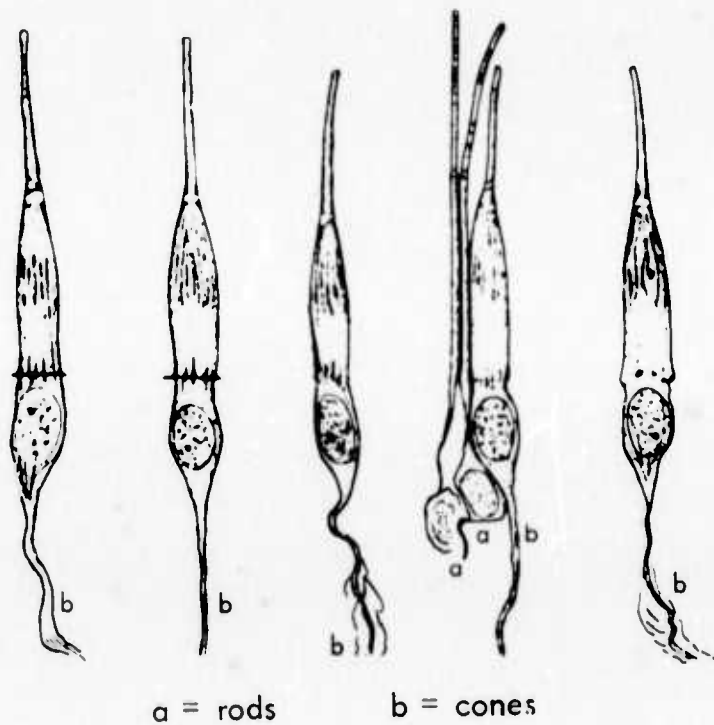


Fig.36—Sketches of rods and cones from an adult Rhesus monkey by Polyak (adapted from Ref. 9)

Enoch made both visual observations and color photographs with retinal preparations illuminated by focused white light (xenon arc) in an apparatus designed for in situ modeling. The cone outer segments showed unequivocal evidence of spectral discrimination that was demonstrably related to the dielectric waveguide properties of the outer segments and their variation over the retina. It is important to note that Enoch's specimens were fully "bleached" in strong white light before his tests.

In Enoch's color photographs,⁽¹⁰⁸⁾ regions in both the central fovea and the periphery show a complete spectrum of cone "types": blue, blue-green, green, yellow, orange, red, etc. The hue distributions in a given region were also found to vary to some degree with angle of incidence of the radiant energy. In many instances this change was the same as that obtained by increasing the wavelength of the incident radiation, another waveguide effect. In his investigations Enoch noted the propagation, singly or in combination, of the HE_{11} , TE_{01} , TM_{01} , HE_{21} , HE_{12} , EH_{11} , and HE_{31} modal patterns.* These results should be compared with those of Snitzer and Osterberg,⁽¹⁰⁷⁾ shown in Fig. 29.

Enoch's results are rich in possibilities and difficult to interpret. But he has, at a minimum, shown definite evidence of the following points:

1. Radiant energy is propagated in human cone outer segments with waveguide modal patterns.
2. Such propagation provides the cones with distinctive spectral separation mechanisms independent of the presence of photopigments.
3. There is evidence of spatial spectral separation within a single receptor outer segment.
4. There is an apparently spatially random distribution over the retina of cone "hue" types, the most populous type being a desaturated yellow-green.

*For terminology, see Ref. 119.

These essentially histological results must be considered in any proposed mechanism for initial photoreception in the human retina.

RECEPTOR DIFFERENTIATION

The concept of a cone photochemical reception mechanism in strict analogy to that in the rods is frequently sustained under the rule of natural parsimony. As has been noted several times in previous sections, the available evidence does not support such a similarity. In the present section some additional physiological data are discussed, indicating that rod and cone mechanisms based on similar photochemical systems are highly unlikely.

Rod/Cone Viability

The work of Noell⁽¹²⁰⁾ on receptor cell viability in the monkey has important implications for differences in human rod/cone mechanisms. Noell found that the intravenous administration of the metabolic poison iodoacetate affected retinal excitation as indicated by the electroretinogram (ERG) in less than 20 seconds after the injection and abolished retinal function in response to illumination within a few minutes. When iodoacetate was administered in doses higher than necessary to block retinal excitation, some of the receptor cell population showed signs of cell death after a few hours. These dead cells disintegrated and disappeared during the following days, but their associated retinal cells--the pigment epithelium, bipolar, and ganglion--survived.

In addition to iodoacetate, irradiation of the retina with X rays also produced photoreceptor cell death in a selective manner. That is, other cells of the retina were unaffected histologically even when a very large fraction of their associated receptor cell population had been destroyed.

Of most particular interest here is the fact that the simian cone cells were far more resistant to iodoacetate and X-radiation poisoning than were the rod cells. The foveal cones appeared normal even when almost all rod cells of the retina had been destroyed. The cone cells outside the fovea did show atrophic changes in their outer and inner

segments, but the cell proper (the nuclear region) survived. Thus the atrophy observed in the outer regions of the peripheral cones was "apparently a consequence of the disappearance of the surrounding rods and resultant lack of mechanical support against the forces pressing the retina toward the pigment epithelium."⁽¹²⁰⁾ The foveal cones were still normal even after nine months of iodoacetate poisoning, which had resulted in widespread disappearance of the rod receptors, including all those close to the central region.

Noell notes that the gross morphological differences between rod and cone cells are unlikely to have any bearing on this phenomenon, since the foveal cones are very similar in form to the rods. Iodoacetate and X-radiation are also likely to be different in their cell action. These results indicate a strong difference in metabolic organization of the rod and cone cells.

Rod/Cone Ontogeny

In ontogenetic studies with the rabbit during the postnatal period, Noell found⁽¹²⁰⁾ that the development of rod and cone cells differed appreciably. At birth, most retinal cells were still in an immature state. Only the ganglion cells were of adult appearance. The bipolars reached adult form only after 7 days. At this stage the receptor cells were very short "buds," but their growth proceeded at a high rate between 11 and 18 days. Their length at about 20 days was 70 percent of adult size. Further growth was slow and closely associated with normal increase in eye size. This, briefly, covers the rod/cone morphology.

However, retinal excitation to illumination, as manifest in the ERG, showed marked differences between rod and cone development not correlated with the above changes in form. The scotopic (i.e., rod-mediated) ERG showed very slow maturation. Its characteristic features appeared at about 10 to 11 days, grew rapidly from the 12th to the 20th day, then continued in steady growth up to 90 days. During the latter period its maximal amplitude doubled and the scotopic threshold decreased by 1 log unit or more.

In contrast to this slow rod maturation, photopic (i.e., cone-mediated) response was well developed at an age of 11 days, had adult form

at 18 days, and changed little after 20 days. Noell⁽¹²⁰⁾ attributed the slow rod ERG development to a steady increase in rhodopsin content during maturation. As evidence for this he noted that the rabbit ERG's at different ages correlated with the ERG abnormalities produced in adults kept on vitamin A deficient diets. The latter produced only scotopic effects.

This evidence seems to strike a fundamental blow to one of the principal arguments advanced by those who support a photochemical receptor mechanism in the cones based on selective photolabile pigments--namely, the argument that since the rod mechanism involves such a photopigment as its basis, the cones should function likewise under the principle of nature's "parsimony" in adapting available techniques. The above data make it seem highly unlikely that the retinal rods and cones could have receptor mechanisms for radiant energy based on similar chemical systems. For if, as the evidence indicated, the slow maturation of the rabbit rod receptors was due to the slow accumulation of the single photopigment rhodopsin to the adult level, then, for a similar photochemical system to be present in the cone receptors, the cones would have to have been supplied with adult concentrations of several photopigments in one-fifth the time required for the rods.

On the other hand, if the cone sensitivity is primarily a result of its physical (i.e., waveguide) characteristics, then the photopic sensitivity development would be expected to parallel the cone morphological development. This is precisely what Noell found.

Rod/Cone Frequency Response

When the human retina is intermittently illuminated under scotopic conditions, the flicker fusion frequency remains constant at about 15 cycles per second (cps). The fusion frequency is independent of retinal illumination over the hundredfold range of 0.1 troland to 10 trolands.⁽¹²¹⁾ Under photopic conditions where the cones are active, the flicker fusion frequency rises to at least 60 cps and is critically dependent on luminance level (the basis of heterochromatic flicker photometry).

It seems unlikely that this rod/cone difference in frequency response could be attributed to differences in bipolar and ganglionic pathways subserving the two receptor systems, since these pathways have much in common anatomically. Rather, the difference is most simply explained by a failure of the rods to respond photochemically at the higher frequencies, with no such failure for the cones. Such a chemical kinetic limitation for the rods seems quite reasonable based on present knowledge of rod photochemistry. But it is difficult to visualize an analogous explanation for the cone behavior, since the variation of photopic fusion frequency with luminance would require a dependence of the kinetics on intensity level.

Brindley⁽¹²²⁾ has, indeed, shown that photopic fusion frequency is not limited by a photochemical mechanism in the cones. He studied the sensations produced by simultaneous stimulation of the human eye with intermittent light and intermittent electric current. He found that lights interrupted at frequencies substantially above that of photopic fusion could still produce visual "beats" with intermittent electrical stimuli transmitted through a conductive eyebath. The results clearly indicated that the "beats" found were those to be expected from interaction of two electrical signals. This implies that the normal photopic fusion frequency cannot possibly be due to a simple photochemical limitation, since the cone receptors were obviously responding to frequencies well above fusion (over twice the fusion frequency).

Again, we are provided with strong evidence against similar photochemical mechanisms in the rods and cones.

RETINAL ELECTROPHYSIOLOGY

Probably the greatest limitation on an adequate understanding of the peripheral mechanisms in human vision is the absence of detailed knowledge of the excitation modes of the retinal neurons. Such information is virtually unavailable for the primate retina, and the data from lower species are as yet poorly understood.

The experimental technique consists in recording electrical signals from microelectrode probes inserted in the retinal or optic nerve

structures of living animals. The type and variation of signals with changes in retinal illumination are then studied. The responses of the ganglion cells are the repetitive spikes typical of neurons in other nervous pathways. But evidence exists indicating that the responses at the receptor and bipolar levels may well be a steady potential difference that varies in magnitude and polarity with the wavelength of retinal illumination. Such a response would violate the so-called law of specific energies. Under this concept, an excited neuron fires in the same manner regardless of how it is excited. But this "law" is only a nonuniversal rule, and there are no data indicating compliance with it in the receptor and bipolar structures. Data on both spike and DC responses are discussed next.

Spike Responses

Granit^(13,123,124) has pioneered the electrophysiological technique of determining overall retinal sensitivity by measuring the relative energy at various wavelengths required to elicit a threshold response of nerve impulses as recorded from microelectrodes inserted into the retinal ganglion cells or their axons in the optic nerve. He found two main types of responses. The most numerous type had a wide spectral sensitivity curve similar to the human luminosity curve, and Granit called this the response of a "dominator" receptor unit. Six or seven other types of responses displayed narrow spectral sensitivity curves located throughout the visible spectrum, and Granit attributed these to "modulator" units. His dominator/modulator theory of color vision^{(70),(75)} is based on these findings.

Optic nerve element recordings such as those of Granit require careful consideration for at least two reasons: (1) in animals that have only rods, the spectral sensitivity curves are broad and show reasonable agreement in shape and maxima location with the absorption curve of rhodopsin; and (2) in animals with cones only (the snake eye, for example), single nerve responses give a broad band curve with a maximum at 560 mμ. The latter type of curve is also observed in mixed rod/cone eyes when the rod activity has been minimized by light adaptation. Hence it may be considered a photopic response.

Examples of such curves are shown in Figs. 37 and 38 for the pigeon, as determined by Donner.⁽¹²⁵⁾ In Fig. 37, Donner's photopic "dominator" response is compared with Granit's earlier measurement. Figure 38 shows the spectral response of a "yellow modulator." It is interesting to note the narrowness of the modulator response, which is a message as transmitted upward through the optic nerve. The response is much more sharply peaked than the individual receptor responses required in a 3- (or 4-) primary receptor theory and is very difficult to explain under such an additive theory.

Granit felt⁽¹³⁾ that these narrow modulator curves probably represented the spectral sensitivities of individual cone receptors, hence providing a mechanism for hue discrimination. He then assumed that a synaptical convergence of several receptors to a bipolar and several bipolars to a ganglion would produce both the dominator (photopic) response and various modulator/modulator or dominator/modulator responses. However, his assumption of synaptic convergence violates the bipolar tripartition results discussed earlier.

Jahn⁽¹²⁶⁾ presented a revision of Granit's interpretation by assuming that the dominator and modulator ganglionic responses were ultimate manifestations of synaptically singular and/or convergent groupings of rods and three types of cones, as illustrated in Fig. 39. Thus Jahn envisioned the dominator response at E_1 as being due to a nerve fiber connected to a group of receptors, while modulator responses at E_2 and E_3 would result from fibers attached to individual cones. On this basis, a modulator response should also be recordable from any single preganglionic fiber, as at E_4 . The luminosity curve was assumed to be determined by the total effect of all dominators and modulators (rather than of the dominator alone, as per Granit⁽¹²³⁾), and hue discrimination would be determined by the modulators alone.

However, the actual divergent/convergent cone-bipolar-ganglion relationship discussed earlier represents a more immediate and anatomically based interpretation of the modulator type responses. Signal processing by means of the bipolar tripartition could result in a multiplicity of modulator type responses at the ganglionic level. Moreover, recordings from ganglion axons are essentially neutral as regards

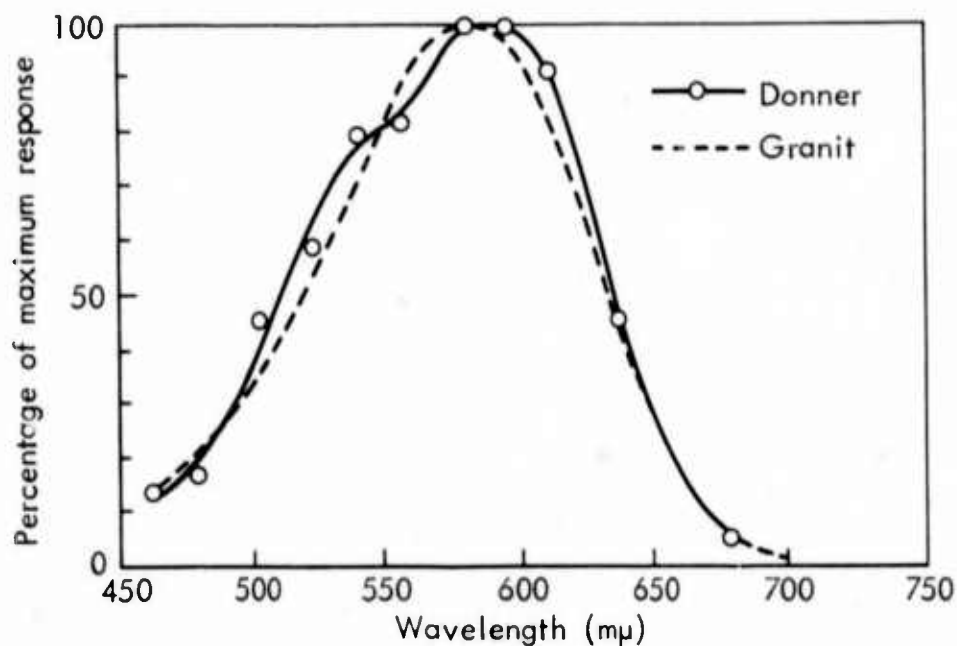


Fig.37—Photopic "dominator" response of the pigeon eye based on the average for three elements by Donner (125), compared with the gross average (multifiber) response determined by Granit (124) (adapted from Ref. 13)

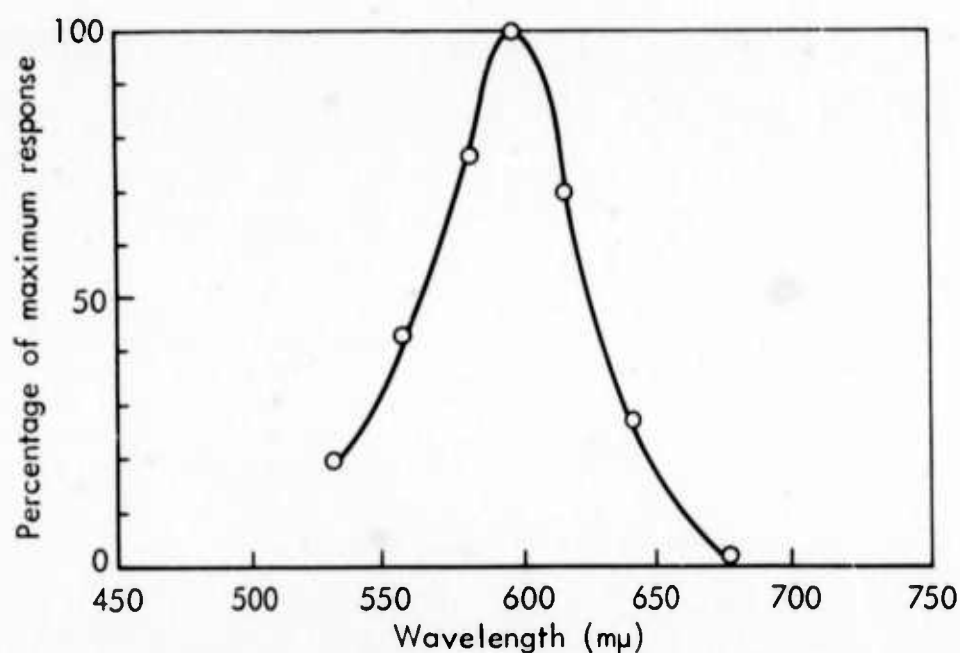


Fig.38—A narrow-band "modulator" spectral sensitivity curve for the pigeon eye by Donner (125), based on the average for three elements (adapted from Ref. 13)

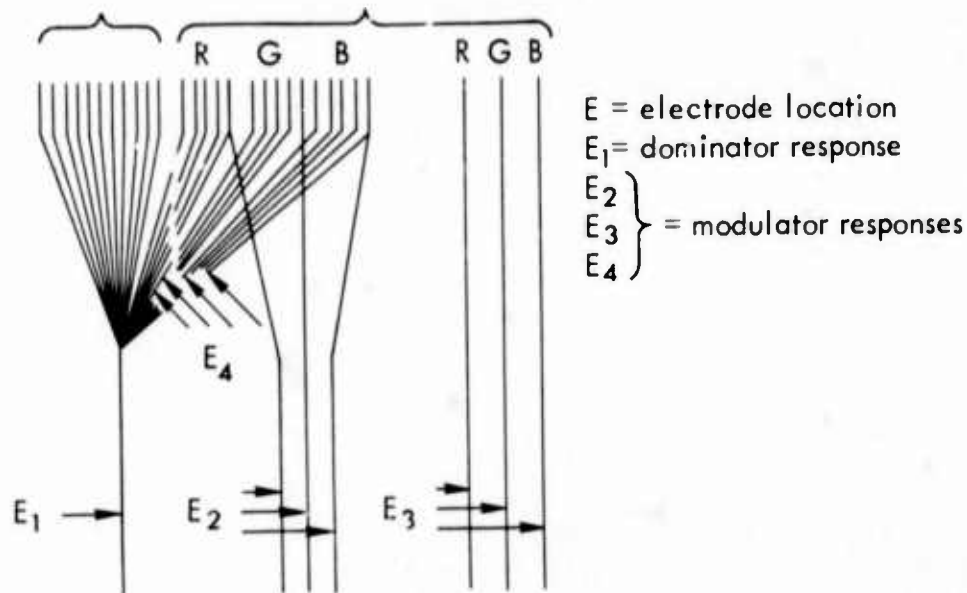


Fig.39 —Retinal receptor systems, according to Jahn, which would explain the electrophysiological results of dominator and modulator curves (Ref. 126)

number of cone types. Evidence of the multiplicity and complexity of single ganglion axon spectral responses that one would expect from a conjugation of spectrally selective cones and bipolar tripartition is offered by the recent data of Chapman.⁽¹²⁷⁾ His extensive recordings of nerve impulses from single ganglion cells of the bullfrog yielded spectral sensitivity curves that in Chapman's own words "suffer an embarrassment of riches, as the varieties of curves do not clearly dictate a few basic color-sensitive mechanisms."

DC Responses

The above ganglion cell or optic nerve fiber recordings are the typical "spike" responses of all-or-none neurons wherein information must be carried by means of a frequency code of some kind. But this frequency code has not been fully deciphered even for the lower vertebrates that are the subjects in the above tests. The most likely

reason for this is that it is still uncertain exactly what information is being encoded. This cannot be determined until the electrical responses of the cone receptors and bipolar neurons are known.

There is experimental evidence that neuronal responses at the cone and bipolar levels are not of the spike type but are, in fact, graded DC electrical potentials. Svaetichin and MacNichol⁽¹²⁸⁾ have reported the results of electrophysiological studies on fish (teleost) retinas, where two fundamentally different types of graded electric response to a light stimulus have been found. Such responses are called "graded photopic responses" (or GPR's) by the authors. Both types start from a negative membrane potential of about 30 millivolts (mv), which is the resting potential level in darkness. The light evoked response is either an increase or a decrease in this negative darkness potential, the new DC level remaining constant for the duration of the light stimulus. The amplitude of the DC potential-change varies roughly linearly with the log intensity of illumination.

On the basis of their reaction to radiation from different parts of the spectrum, the GPR's have been separated into achromatic (L) and chromatic (R-G and Y-B) types. Using automated procedures, oscillograms of the spectral response curves were obtained by subjecting the retina to 24 different spectral stimuli of equal energy ranging in succession from 400 to 750 mμ. The types of results are illustrated in Fig. 40. The L type of response (luminosity response) was always found to be an increase of the negative resting potential, i.e., hyperpolarization, independent of the wavelength of the light stimulus. The peak of the L spectral response curve was between 500 and 600 mμ. The curve was smooth in some species, and had submaxima in others. It was found in all species studied and was the only type found in fish caught at depths of 30 to 70 m, which presumably do not need hue discrimination.

The chromaticity responses (R-G and Y-B) are characterized by having two maxima of opposite polarity in their spectral sensitivity curves, as shown in Fig. 40. These responses peaked either in the red and green (R-G responses) or in the yellow and blue (Y-B responses) regions of the spectrum. In these chromaticity responses there was always a neutral point in the spectrum, with shorter wavelengths giving

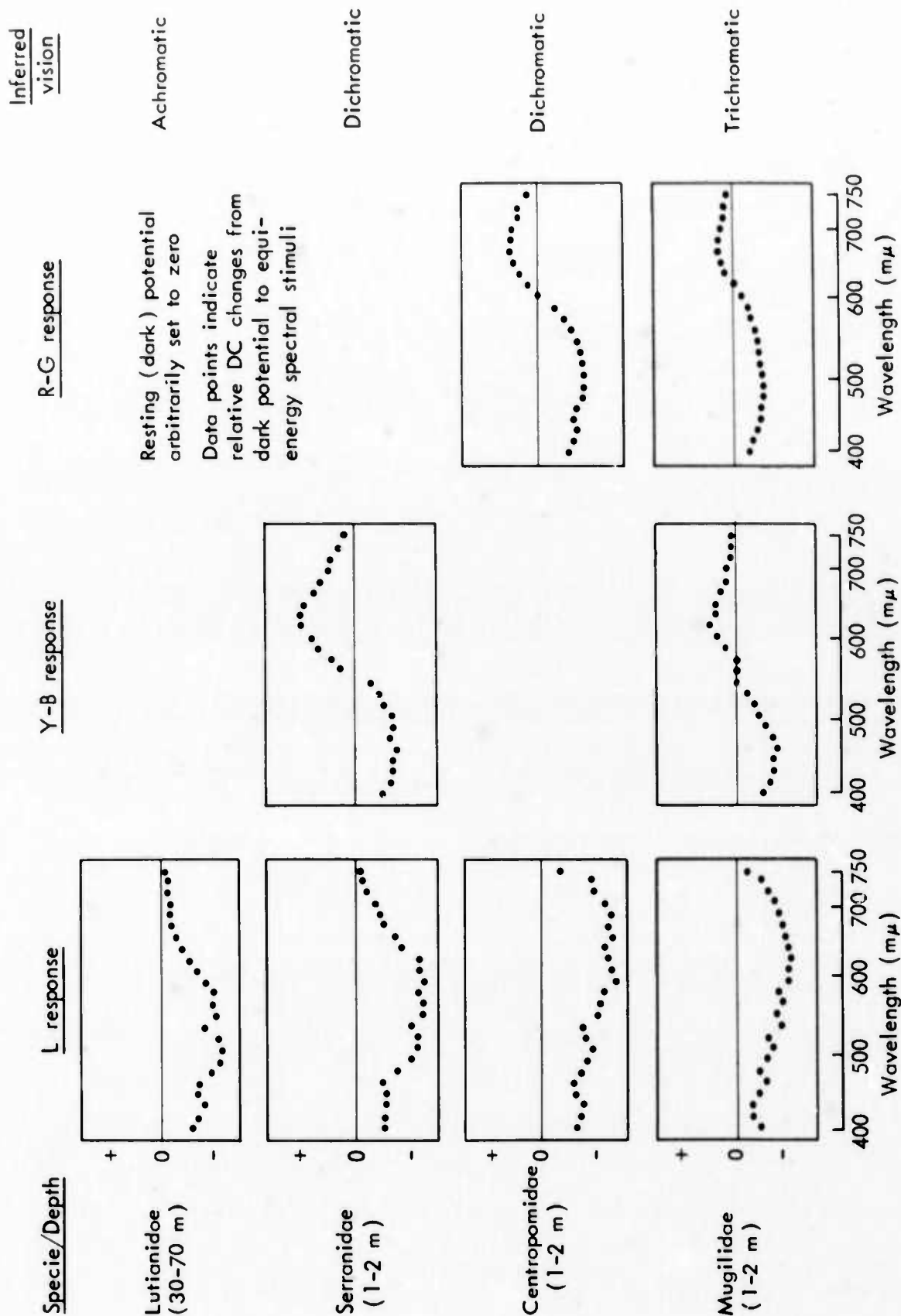


Fig.40—DC electrical responses to monochromatic stimuli in retinal neurons of fish (teleost)
(adapted from Ref. 128)

generally negative response and longer wavelengths giving generally positive response.

Further GPR measurements were made with a selective adaptation technique using a steady monochromatic background illumination of various wavelengths with a 0.3-second test flash. These results led Svaetichin and MacNichol to conclude⁽¹²⁸⁾ that the R-G and Y-B chromaticity responses were each algebraic resultants of two separate potentials of opposite sign. This result, together with measurements of electrode position, indicated that the Y-B and R-G responses were from the bipolar cell layer. On the other hand, the shape of the L curve remained constant regardless of the wavelength of the background illumination. The only "adaptation" effect on the L response was a reduction of sensitivity at all wavelengths. This result, corroborated again by probe location, led Svaetichin and MacNichol to believe that the L response was obtained from the synaptic endings of cones and horizontal cells.

Whatever the origin of the GPR results, their possible implications must be considered.* If their information content is assumed to be frequency encoded in the ganglion cell discharges, they could satisfy the basic requirements for photopic (i.e., color) vision. It is interesting to note that very recent work by Rosenberg, Heck, and Aziz⁽¹²⁹⁾ has shown that diphasic responses such as the Y-B and R-G responses of Fig. 40 can occur in a single biased organic cell under illumination. They found Svaetichin type responses from a β -carotene cell as a result of simultaneous, opposing photovoltaic and photoconductive effects. The total response curve had a unique shape for each region of the visible spectrum.

None of the Svaetichin type electrophysiological measurements have yet been made on primates, but the teleost results still present intriguing possibilities as regards color vision in man. The above conclusions as to the origin of the DC responses are, of course, tentative. But since ganglionic responses are known to be of the spike type, the

*Some possible implications, with a summary of additional work, are discussed in Ref. 130.

GPR signals must be confined to the more peripheral neurons. If the bipolars have only DC response characteristics, then the visual system has an analog link and is no longer purely binary in nature. This would specifically deny the applicability of the law of specific energies to the bipolars and remove one of the principal objections to a Hering opponent-colors type theory of color vision. One could in fact devise a "continuum" peripheral mechanism wherein the responses of a multiplicity of cone receptor types are analyzed by the three bipolar types to result in biasing signals governing the spike firings of the ganglion cells.

CONCLUSIONS

Some of the more important conclusions reached in this discussion of peripheral physiology may be summarized as follows:

1. There is still no definitive experimental evidence for the existence of three specific spectrally selective photopigments in human cones for the mediation of color vision. Metamerism continues to be the principal support of the view that such pigments do exist.
2. In the central area of the human retina, where color vision is fully developed, each cone has directly associated with it three bipolars. This ratio decreases to one bipolar per cone in the periphery.
3. There is good correlation between the bipolar/cone ratio and the color zones of the visual field. A ratio of 3:1 gives full color vision, 2:1 gives reduced (dichromatic) color vision, and 1:1 gives achromacy.
4. Radiant energy is propagated in human cone outer segments with waveguide modal patterns. Coupled with a variation in outer segment diameters and indices of refraction, such a modal propagation provides the cones with a spectral discrimination mechanism.
5. There is apparently a spatially random distribution over the central retina of cone "hue" types based on the waveguide

mechanism, with the most populous type having a dominant wavelength in the range from 560 to 580 mμ.

6. There is strong evidence indicating that the rod and cone receptors have quite different metabolic organizations and photochemical systems.
7. Evidence from lower vertebrates indicates the distinct possibility that electrical responses at the bipolar level are graded DC potentials rather than the spike discharges of the ganglion cells.

Combined with the results of Section IV, the definite implication of the above conclusions is that the mechanism for human color vision is not based simply on three cone photopigments, but rather on a combination of variation in cone physical characteristics (including any pigments) and multiple bipolar "analyses." This in turn implies that the Young-Helmholtz three-pigment theory, the Hering opponent-colors theory, and the various "stage" or "zone" theories may each contain some element of truth, but that none is individually capable of providing a suitable explanation of the human visual mechanism.

VI. CEREBRAL PHYSIOLOGY

A crucial and as yet unresolved question in human color vision is, What aspects of the overall process have their physiological basis in the retina and what aspects are resident in the brain proper? The tendency has been to assign as much as possible to the peripheral structures and then categorize the balance as central phenomena. Such an approach results in a deemphasis of the highly elaborate cerebral structures in the visual pathways of primates and may very well prove to be a false "economy."

Neurohistologists have long noted the morphological similarity in the peripheral portion (i.e., the retina) of the visual system in all the vertebrates. The same structural principle, namely, a receptor layer followed by two consecutive neuron links, is in general preserved from fish on up to man. Unlike the retina, however, the central nervous apparatus of the optic system exhibits, in vertebrate evolution, essential and major changes, along with its increasing organizational complexity.⁽¹³¹⁾ This is illustrated in Fig. 41. In the visual system of the lowest vertebrates, the fish, the optic nerve fibers from the retina terminate directly in the midbrain tectum (superior colliculus), where there is little or no neuronal differentiation into layers. In higher vertebrates, such as birds, there appears a relay system, the thalamus, for a portion of the afferent fibers from the retina. The balance still proceed directly to the midbrain. It is only at this stage (i.e., after development of a subcortical relay synapse) that an evoked potential appears in the just-developing forebrain in response to light incident on the retina.

The importance of this relay system can be judged by the fact that in the highest vertebrates, the mammals, almost all of the optic nerve fibers synapse in a subcortical relay system. A new system of nerve fibers then proceeds to well-developed higher brain centers. This relay network is highly developed in man. As indicated in Fig. 3, virtually all the optic nerve fibers synapse within the lateral geniculate nuclei of the thalamus. Only a relative few terminate at the other subcortical centers, which have the dominant role in lower vertebrates.

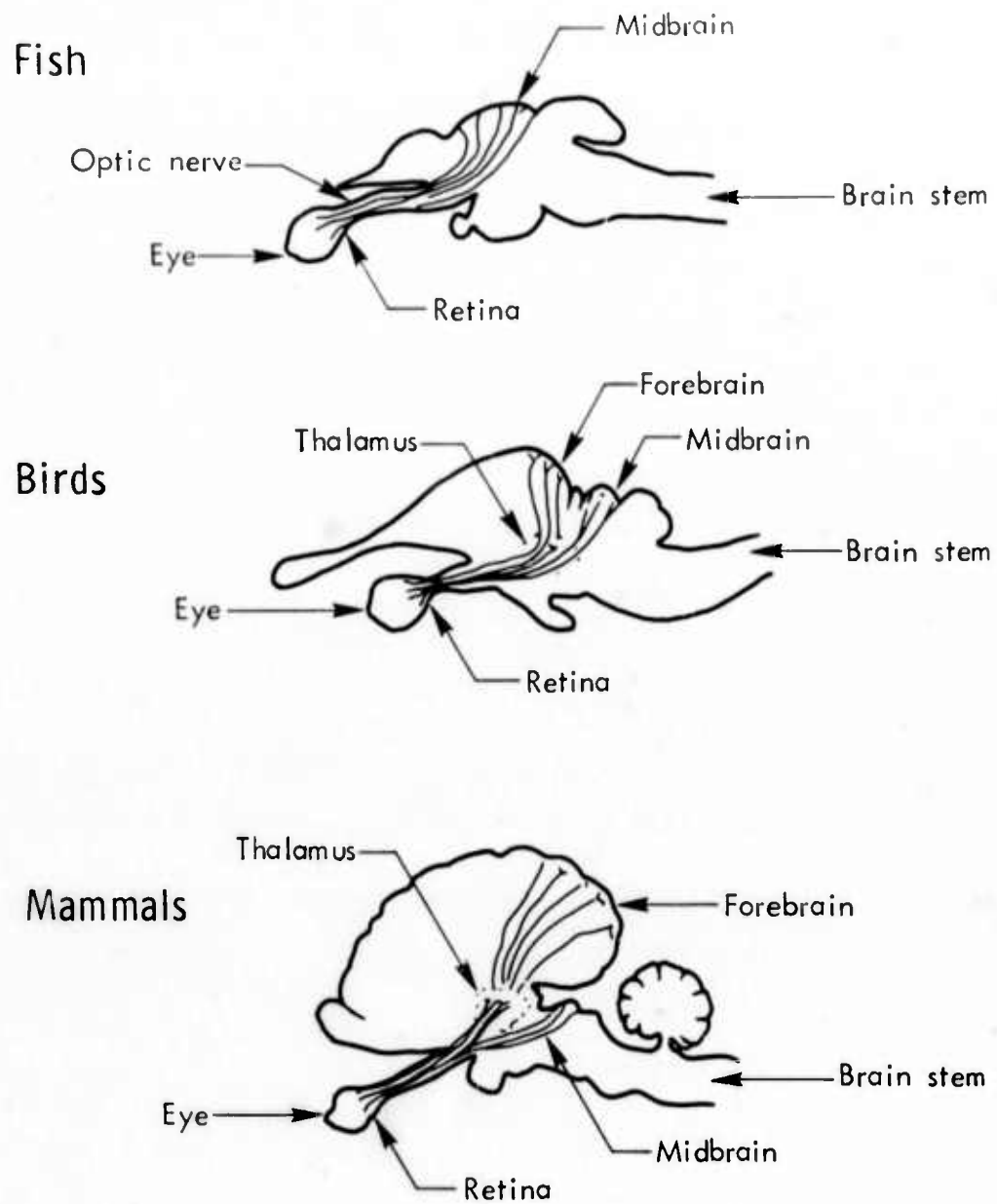


Fig.41 —Evolution of the central visual system in vertebrates

For the latter, for example, the pregeniculate nucleus appears to serve as a central afferent link of the pupillary reflex pathway.⁽⁹⁾ In all mammals, from the lowest to the highest, part of the retinal fibers continue directly to the midbrain tectum, but this direct connection is reduced in the primates to a very few fibers. The dominant role in the subcortical visual system is here taken by the lateral geniculate nuclei.⁽⁹⁾

SUBCORTICAL SYNAPSES

In man, after partial decussation in the chiasma (see Fig. 3), the crossed (contralateral) and uncrossed (ipsilateral) optic nerve fibers are completely segregated in the lateral geniculate nuclei, where they terminate in separate layers.⁽⁹⁾ The LGN is made up of six well-defined cell laminae that are quite separate in the area subserving central vision but are partly fused with each other in the area subserving peripheral vision.

The structure and inferred synapses in the LGN are shown in Fig. 42, adapted from Polyak.⁽⁹⁾ The foveal area of each nucleus is divided into contralateral (1c, 2c α , 2c β) and ipsilateral (1i, 2i α , 2i β) sets of three layers each, one set for each eye. The near peripheral areas contain only two layers each (1c, 2c α ; 1i, 2i α) and the far peripheral areas only one, the latter contralaterally connected. From the LGN, nerve fibers (VR) proceed to the right and left striate area of the cortex.

In the primate LGN, the first and second layers, (1c) for the contralateral eye and (1i) for the ipsilateral eye, consist of large cells, while the other four layers (2i α and 2i β , 2c α and 2c β) contain small cells. The cells of all six layers have neuronal projections onto the visual cortex (see Fig. 3). If one eye is removed surgically, three layers in each half of the LGN undergo rapid degeneration: the first (1c), the fourth (2c β), and the sixth (2c α) of the opposite geniculate nucleus, and the second (1i), the third (2i α), and the fifth (2i β) of the ipsilateral nucleus.^{(9),(132)} Thus each nucleus consists fundamentally of two sets of three layers each, each set related to one eye.

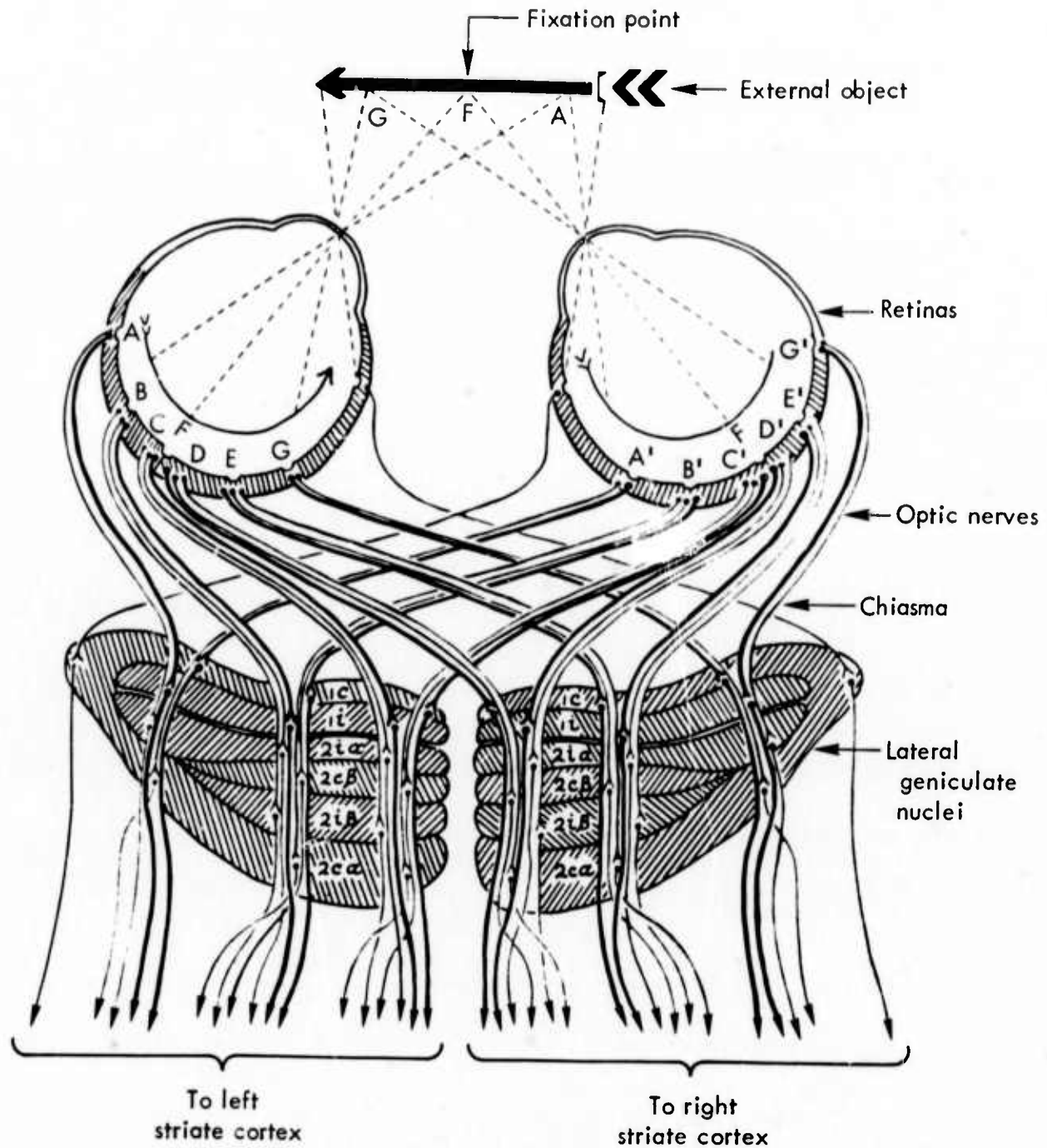


Fig.42 — Schematic of the primate visual system emphasizing the structure and inferred synapses in the lateral geniculate nucleus (adapted from Ref. 9)

If a very small and localized lesion is made in the central retina of one eye, a spot of cell atrophy appears in all three of the corresponding LGN layers.^{(9), (132)} The retinotopical organization of the lateral geniculate nuclei is thus well differentiated.

Furthermore, this organization is retained in the supranuclear portion of the visual pathway, the striate area of the occipital lobe (see Fig. 3). If small lesions in the shape of various "figures" are made in the striate area, the zones of degeneration in the lateral geniculate nuclei agree rather accurately with the lesions in shape and corresponding location.⁽⁹⁾ In short, there seems little doubt as to the essential retinotopical organization of the entire visual system--retina: lateral geniculate nuclei: striate cortex. Still unknown, however, are the detailed structure and interconnections of this system.

Polyak postulated⁽⁹⁾ that each optic tract is composed of two or three categories of optic nerve fibers, since each tract is related to two layers in the peripheral segment of the lateral geniculate nucleus and to three layers in the foveal segment of that body. This is the reasoning behind his synaptical schema in Fig. 42. This implies that from each local spot in the central area of the retina, three types of nerve fibers (and in the periphery two types) proceed up through the optic tract, one to each lamina of the LGN. An alternative is for each optic nerve fiber to divide into three branches, one going to each layer on reaching the LGN, but intensive study has failed to show any evidence of this.⁽¹³²⁾

The bulk of the cellular elements of the laminated lateral geniculate nucleus in humans appears to be typical multipolar neurons rather than the highly specialized bipolar form.⁽⁹⁾ To each of these "principal" cells is synapsed a single terminal nodule from an arborized optic nerve fiber. Each fiber synapses, through the arborization, to about six LGN cells in the given layer, and the nodules are attached to the cell bodies (axosomatic synapse). Thus each optic nerve fiber synapses with several LGN multipolar neurons, but each of the latter receives only a single optic nerve synapse. There seems to be no overlap.⁽⁹⁾ The dendrites of the LGN cell, on the other hand, radiate out in all directions and are approximately of the same length, so that there is

a reciprocal overlapping of the dendrites of adjacent LGN cells amounting to about half their spread. A single axon emerges directly from the LGN cell body and presumably heads for the cortex, but it is soon lost in the surrounding fibers so that it cannot be traced.

In addition to the "neighbor" interaction, the principal LGN nerve cells are further interconnected by association cells that synaptically relate approximately a dozen nearby principal cells. Polyak⁽⁹⁾ suggests that these cells may serve to "condition" a group of principal LGN cells, possibly under the influence of the striate area of the brain proper. This may be inferred from the presence of corticofugal nerve fibers, probably from the striate area, terminating in an arborization that spreads over an area occupied by several hundred principal nerve cells in a given lamina.

Polyak⁽⁹⁾ gives considerable admittedly inferred elaboration on the above and other synaptical relationships in the various layers of the lateral geniculate nucleus, but these details show no clear correlation with color perception. He does note, however, that the laminar organization of the lateral geniculate body in man appears essential for color perception, and this body is not just a simple relay between the retina and the brain proper. Binocular color mixing,⁽¹³³⁾ for example, where color pairs neither completely fail to fuse, nor fuse perfectly⁽¹³⁴⁾ as in one eye alone, quite probably has at least partial explanation in the LGN. It should also be noted that there is sufficient experimental evidence⁽⁹⁾ to conclude that in human adults there is considerable individual variation in the pattern and complexity of the arrangement of the cell and fiber layers in the lateral geniculate nuclei.

From both topographical studies and available pathological evidence, Clark⁽¹³²⁾ has divided the human LGN into zones according to which retinal area is subserved. This is illustrated in Fig. 43, which shows a transverse section through the human LGN, with only the laminae related to one eye being represented. The nomenclature is analogous to that of Fig. 42.

Figure 43 indicates several features of the nerve synapses in the layers of the LGN. As the fixation point is approached, the relative

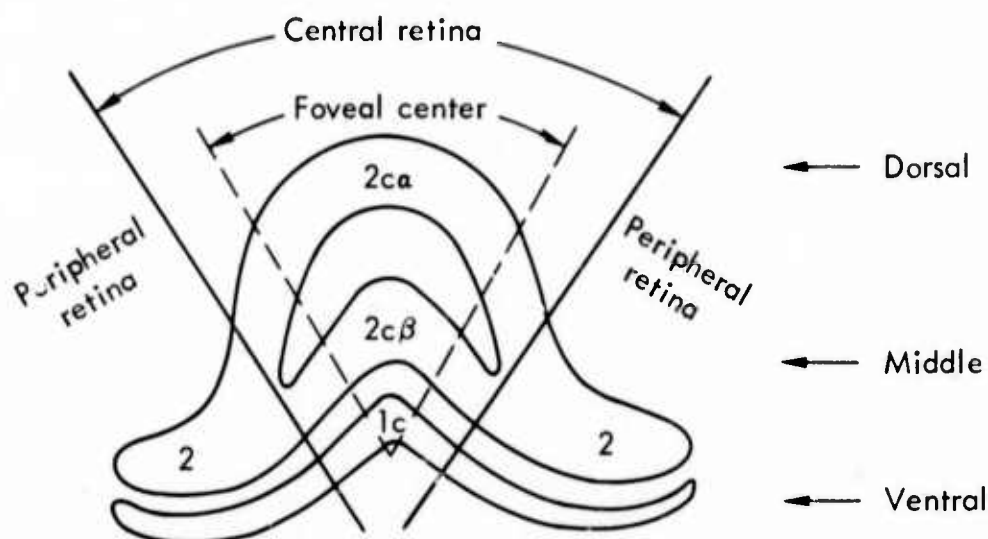


Fig.43 —Zones of the human lateral geniculate nucleus
(adapted from Ref. 132)

number of LGN cells in layer 1 subserving a small retinal area is reduced. For the very center of the fovea, optic nerve synapses in layer 1 appear to be absent altogether. On the other hand, as the peripheral retina is approached, the relative number of layer 1 LGN cells subserving a retinal area increases. This evidence strongly indicates that layer 1 subserves a diffuse receptor system, perhaps via the mop bipolar and diffuse ganglion neuronal links. The nondifferentiated portion of layer 2 would thus serve as the terminal synapsing area for the mid-ganglion subserving a localized receptor system involving both rods and cones in the periphery.

By extension of this reasoning, the whole of the No. 2 laminated portions of the LGN may be considered to subserve synaptically very small retinal areas and primarily the retinal cone system through the latter's related bipolar and ganglion cells. This view is also indicated by the apparent evolutionary development of the LGN suggested by

Polyak's schematic arrangement in Fig. 42. The contralateral and ipsilateral laminae 1 are clearly separated from the whole of laminae 2 and further contain histologically different cells from those in laminae 2. From its inner back-to-back contralateral-ipsilateral 2β layers separating the two 2α layers in the central retinal zones, the LGN appears to have developed and differentiated its laminae 2 for a function basically independent of that for which the laminae 1 evolved.

Apparently the most direct partial experimental verification of these views would be to observe the LGN atrophy resulting from an exceedingly small experimental lesion carefully made in the very center of the rodless central fovea of the primate eye, an extremely difficult task. Clark⁽¹³²⁾ discusses joint pathological/histological studies by Ronne that support the view that LGN laminae 1 are primarily rod subservient, and laminae 2 are cone subservient.

If in laminae 2 it is assumed (as histological evidence suggests^{(9),(132)}) that the ratio of optic nerve fibers to LGN cells is sensibly constant, then the number of ganglion axons synapsing in lamina $2c\alpha$ from a given central retinal area outnumber the synapses in lamina $2c\beta$ from the same retinal area. This disparity apparently increases right up to the very center of the fovea. As a cone receptor always has an associated midget-bipolar/midget-ganglion neuronal link, this provides strong evidence that the midget ganglions have their terminal synapses in laminae $2c\alpha$. This disparity might further imply that any hue discriminations mediated by laminae $2c\beta$ would be less specific than those mediated by laminae $2c\alpha$.

The conjectures of the preceding paragraphs that the three layers of the LGN subserving one eye are the sites of synapses of three distinct types of ganglionic activity have considerable confirmation in the recent work of DeValois, et al.^(110,135,136) In his schematic of Fig. 42, Polyak shows three types of ganglion axons from each local spot of the central retina proceeding to three corresponding layers of the LGN. Polyak called these the "red, green and blue fibers" to correspond with the trichromatic theory. With extensive recordings from microelectrodes in simian LGN, DeValois has found three distinct types

of LGN cell response carefully segregated in the three layers corresponding to the appropriate eye. But the segregation is on the basis of response mode and not of the region of the spectrum to which a cell responds. He found cells in all three layers responsive to light throughout the visible spectrum.

DeValois⁽¹¹²⁾ found that the LGN cells were ON, OFF, or ON-or-OFF types, based on their firing activity at the onset and termination of a visual stimulus. In the dorsal layers ($2c\alpha$ and $2i\beta$), DeValois found predominately ON-cells that fired with a sustained burst of activity at the onset of the stimulus and lasted for the duration of the stimulus. Further, virtually all the dorsal LGN cells subserving the foveal region of the retina had narrow spectral sensitivity curves. The average response curves for the five types of ON-cells found are shown in Fig. 44. The narrow 550 and 590 peaked curves have no secondary maxima, but the 620 peaked "red" curve has a slight secondary peak in the far blue. All 440 peaked "blue" cells found by DeValois had a secondary peak near 505 m μ . Most of the 505 peaked cells in turn had a secondary peak at 440, but some were found with secondary maxima at 550, 590, or 620. The narrow 550, 590, and 620 cells predominated in the foveal region, and the sensitivity of these curves remained unchanged from light to complete dark adaptation. On the other hand, in the peripheral regions many of the LGN ON-cells had more than one peak, and the peak sensitivity changed from the light to dark adapted state.

In the middle layers ($2c\beta$ and $2i\alpha$), DeValois found ON-or-OFF cells. Here, for a given wavelength a cell would fire either at the onset or at the termination of a visual stimulus, but not at both. A given firing would be accompanied by an inhibition in the opposite sense from its spontaneous firing level. Two distinct types of these cells were found, a red-green and a yellow-blue. The response of a single "red-ON, green-OFF" cell to equal energy monochromatic stimuli is shown in Fig. 45. The similarity of these responses to the Svaetichin GPR responses in fish is obvious. The ON-or-OFF cells gave no response to white light even at intense levels. This strongly suggests that the

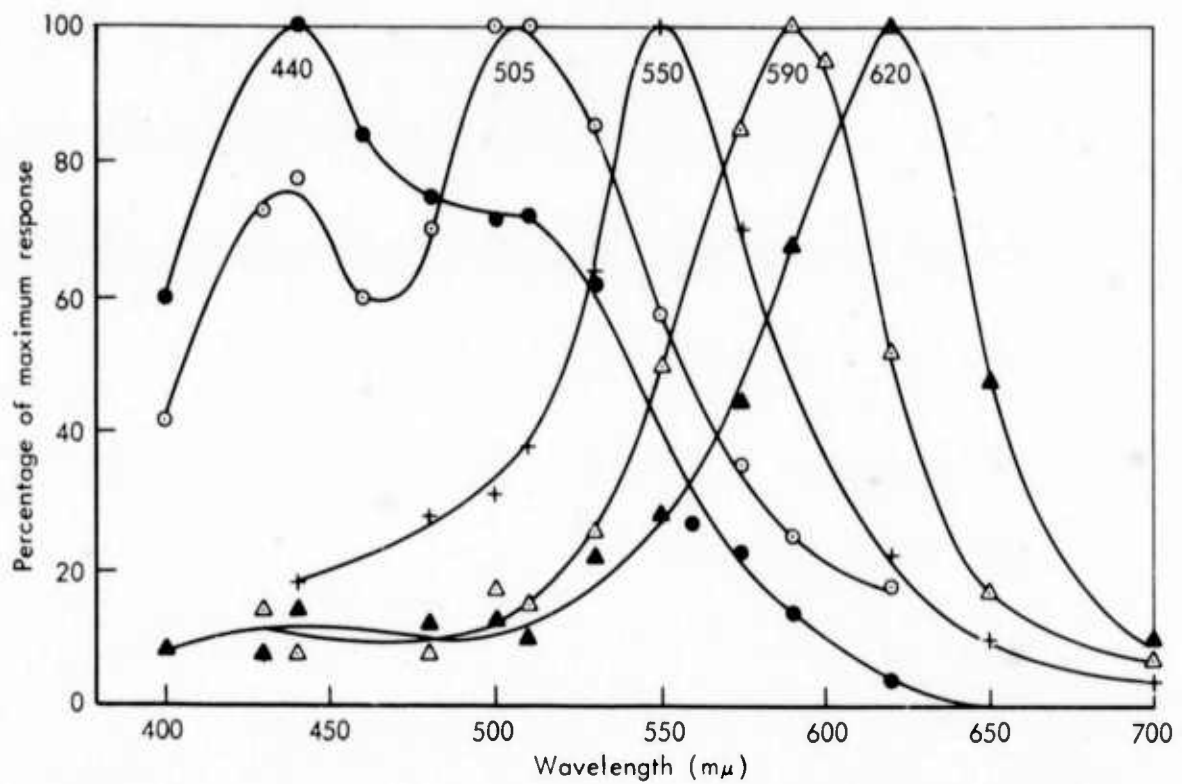


Fig.44—Average lateral geniculate nucleus ON-cell response curves (adapted from Ref. 110)

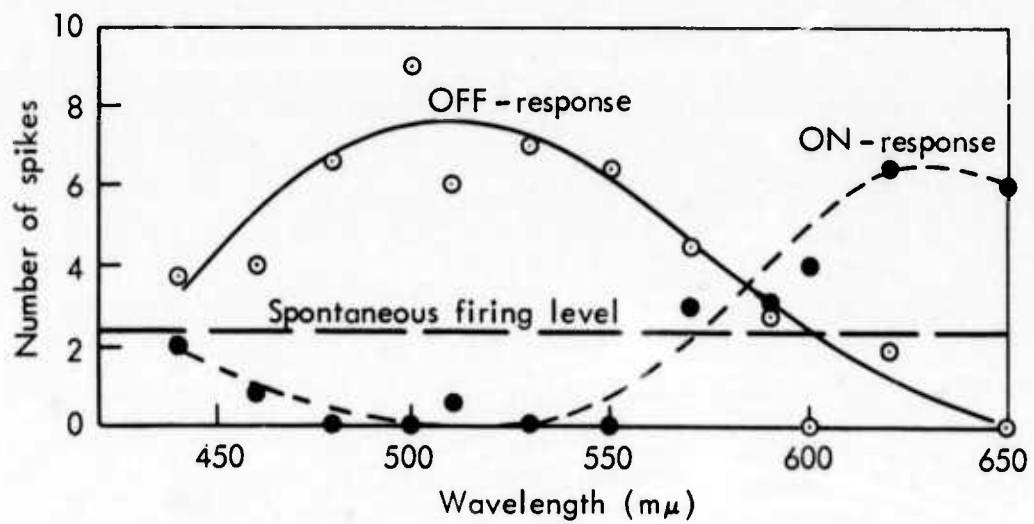


Fig.45—Response of a single "red-ON, green-OFF" cell in middle layer of the lateral geniculate nucleus (adapted from Ref. 110)

middle layers of the LGN are involved in phenomena of color adaptation, particularly in the phenomenon of color constancy with changes in illuminant color.

The studies of the most ventral layers (li and lc) were not as extensive, but DeValois found only OFF-cells there. These cells exhibited a high spontaneous firing rate that was inhibited at the onset of the stimulus, followed by a firing burst at offset. Few cells were found with narrow spectral sensitivity curves, most having broad response curves with a peak at 505 or 550 m μ or both. This correspondence with the peaks of the overall scotopic and photopic luminosity functions indicates the possibility of brightness mediation by the ventral layers of the LGN. Smith, DeValois, and Kitai noted⁽¹³⁷⁾ that the inhibitory phase of the OFF-cell response was very sensitive, a 505 peaked OFF-cell responding to a stimulus 2 log units or more below that required to elicit a response from a 505 peaked ON-cell in the dorsal layers.

In summary, the above results by DeValois, et al., flatly refute attempts such as those by Clark⁽¹³²⁾ and Talbot⁽¹⁸⁾ to involve the three LGN layers directly in a 3-receptor system. Hartridge⁽¹³⁸⁾ has also attacked Clark's postulates on other grounds. Of course none of these latter authors had the benefit of the more recent findings that unquestionably show a dynamic role of the LGN in primate color vision. But this role is far more complex than a simple sorting of three primary signals. The anatomical riches of this body shout for active inclusion in any proposed mechanism for color vision. They also demand caution when the retinal details of lower vertebrates are compared with those of man. One must remember that man has a fully developed lateral geniculate nucleus and can process much more information from the peripheral structures than can the lower vertebrates. For example, the chicken may very well require its color filters in its cones, despite their iodopsin content, because of its relatively undeveloped precortical structures.

VISUAL CORTEX

The experimental evidence is conclusive⁽⁹⁾ for placing the final termination of the afferent signals of the human visual system in the occipital lobes of the cortex. This region is known as the "striate area." The visual afferent axons terminating here originate in the lateral geniculate nuclei.

The retinotopical organization of the striate area, i.e., geographical projection of the retina on the cortex, has also been confirmed. Each area of the retina is fairly sharply represented in a corresponding area of the visual cortex. This is illustrated in Fig. 46. On the left, the striate area of the hemisphere of the brain is shown. On the right is the right half of one visual field. The shadings indicate the representations of different segments of the visual field on the cortex.

While this point-to-point retina/cortex correspondence appears from clinical studies⁽¹³⁹⁾ to be rigid and permanent, equally extensive observations indicate that the organization of function in the visual cortex is not rigid. Activity there is not determined exclusively and inevitably by impulses that reach each spot in it from corresponding points in the two retinas.⁽¹³⁹⁾ Further, there is no evidence of triple mode response in the striate cortex analogous to that in the LGN.⁽¹³⁷⁾

The function of the striate area appears to be under the control of portions of the cortex immediately adjacent to it. This is indicated by clinical cases involving disease or injury in the immediate neighborhood of the visual cortex. For example,⁽¹³⁹⁾ subsequent to a slight stroke, a highly educated man was unable to read. He could not identify letters of the alphabet by sight or name, and could recognize them only by tracing. He could write correctly, though he could not read what he had just written. Although he could "see" the letters, he could not recognize them. The specificity of this defect is indicated by the fact that his identification by sight of other objects, including numerals, was undisturbed. His use of arithmetic did not suffer!

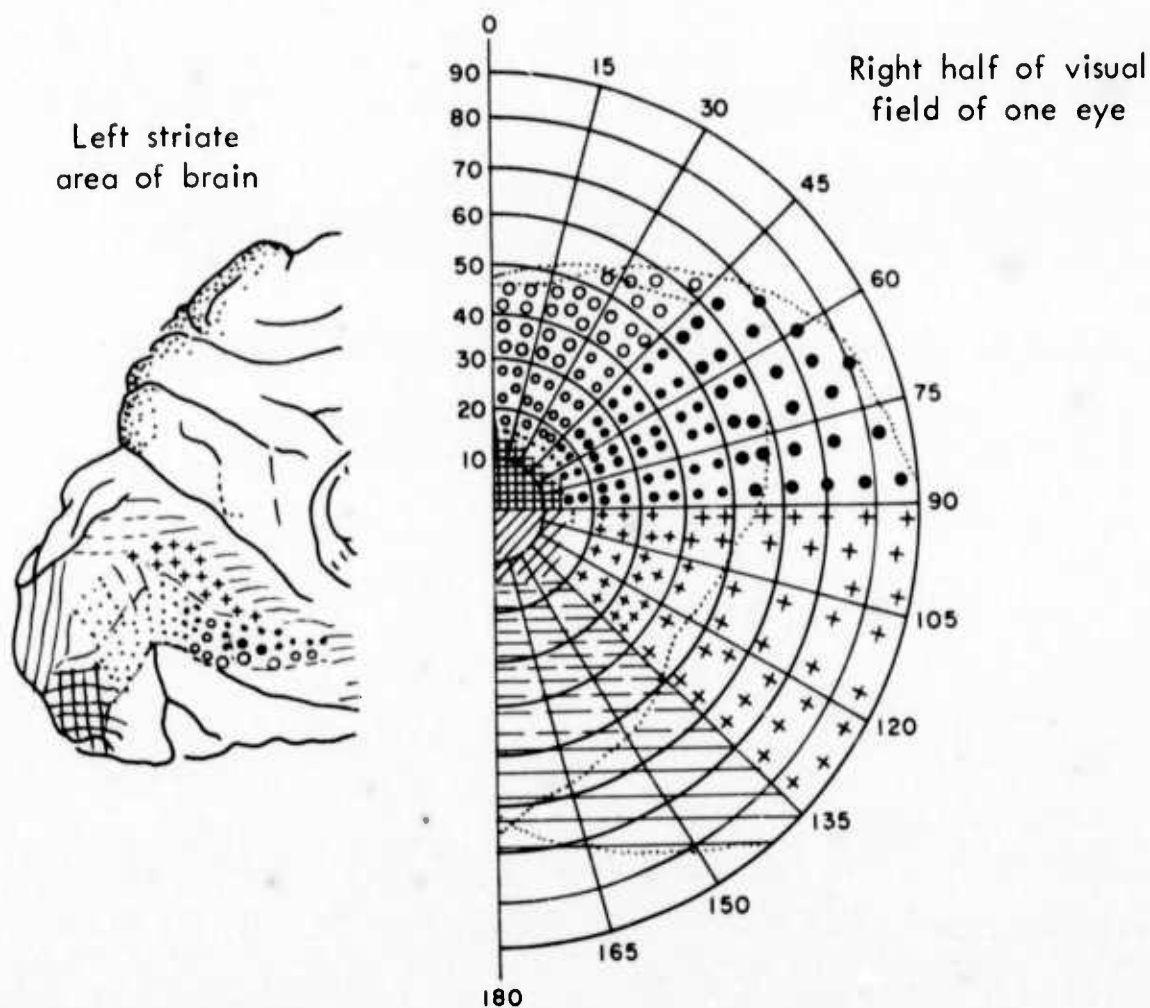


Fig.46—Retinotopical organization of striate cortex (adapted from Ref. 139; permission granted by C. V. Mosby)

In another case, ⁽¹³⁹⁾ a professional artist was unable, subsequent to a stroke, to use colors, stating that they had lost their significance to him. He was not color-blind, and in tests could identify the principal hues. But he had lost the ability to discriminate between gradations of a single hue (various shades of red, for example) and was unable to associate colors with familiar objects.

Direct evidence of control by the higher brain centers on the visual cortex has also been shown by noninjury procedures. Haider, Spong, and Lindsley ^{(140), (141)} recorded via the electroencephalograph the

potentials evoked from the human cortex by visual stimuli. The subjects experienced a vigilance task requiring a visual discrimination and a mechanical response. Dim "response" flashes were aperiodically interspersed among brighter "no-response" flashes. They found a direct correlation between the subjects' attentiveness and the magnitude of the cortical response evoked by detected stimuli. Similar studies by Davis⁽¹⁴²⁾ with auditory stimuli also indicate that requiring the subject to make a rather difficult sensory discrimination results in an enhancement of the cortical potentials evoked by discriminated stimuli.

More detailed evidence of the functioning of the visual cortex and its relation to the higher associational centers of the brain is notably meager, particularly with reference to those aspects concerned with color perception. The general conclusions that can be drawn from observations on human subjects are as follows:

1. Primary visual perception, including color vision, perception of form, and relative localization in space is subserved by the striate area of the cortex.
2. There is geometric projection of the retina on the striate area.
3. The functional organization of the striate area is not rigidly determined by this point-to-point representation, but is dependent on, and modifiable by, even higher centers of the brain.

VII. PSYCHOPHYSIOLOGY

Direct experimental results are not available to indicate the physiological mechanisms involved in "processing" the suprarretinal visual information to yield sensory perception. But there is an abundance of indirect data that may contain clues to such underlying mechanisms. These studies are primarily concerned with visual sensations, particularly under differing conditions of observation, and may be considered principally in the realm of psychology. Emerging correlates should be considered as psychophysiological rather than psychophysical, since the experimental procedures inherently require that the subject describe (by verbal or other response) what he perceives. Therein lies a problem confronting the investigator: The subject constitutes an absolute majority of one in any assessment of the "validity" of his description as a meaningful indication of his sensation. There are, of course, many different procedures that collectively tend to obviate this problem, but such measurements cannot be considered objective until the relevant physiological variables are identified and can be specified, if not actually controlled.

Despite this absolute individuality with respect to a given subject's sensations, some fairly general correlations have been established between specified external viewing conditions and what a "normal" observer perceives thereunder. In a broad sense these correlations may be considered as manifestations of the organism's attempt, by means of physiological processes, to "adapt" to the encountered conditions in such a way as to maximize the "normality" of the perceived sensations. In other words, there is a striving for constancy in the perception of familiar things. This phenomenon is therefore known as adaptation.

ADAPTATION

There is an enormous amount of literature on the role of adaptation in human vision. The following brief discussion of some aspects of this phenomenon is included to emphasize that the human visual mechanism does not receive its primary explication in retinal structures

alone. Some effects of adaptation on color-matching and luminosity results have already been discussed in Section IV. The present section is intended as a qualitative overview. Much more detailed discussion may be found in Evans,⁽³⁾ on which the present review leans heavily.

Brightness Adaptation

On entering a dimly lit room from bright sunlight, a few minutes are required for the eye to adapt to the lowered level of illumination. After complete adaptation, however, visual sensations of brightness in ordinary scenes seem the same, despite the wide divergence in the illumination levels. Evidence indicates that when the eye is adapted to any photopic luminance L , the sensation of luminosity does not vary with L . The psychological brightness sensation is independent of luminance if adaptation to this luminance level is complete. For reasonable levels of illumination, one's room seems about as well lighted by artificial light as by sunlight.

This phenomenon is known as brightness adaptation, and its enormous dynamic range is not explicable in terms of the variation in pupillary diameter. It probably involves both retinal and cortical adaptive processes, and the possible role of the ventral layers of the LGN was mentioned in Section VI. As was illustrated in Fig. 7, the range of working luminance levels in photopic vision is about one million to one. The visual system accomplishes this large variation by a change in its overall sensitivity. The total range it can appreciate simultaneously is much smaller, being of the order of 10 to 1 at the lowest levels and only about 1000 to 1 at the best possible levels in full daylight.⁽³⁾

In a general way, the sensitivity level of the visual system is determined by the total light entering the eyes. For any given set of conditions, the current sensitivity level is referred to as the brightness-adaptation level. If this level corresponds to the luminance L_a , then there is a lower luminance L_b at and below which all stimuli, regardless of their spectral composition, will be perceived

as black. This level is usually called the "black point," and it rises and falls with the brightness adaptation level. It is a complex function of both the adaptation level itself and the stimulus that is perceived as "white."

The situation is illustrated in Fig. 47, where the long solid curve represents the range of intensities to which the visual system may adapt itself. When adapted to a given level, say L_a , the short intersecting curve represents the range of intensities corresponding to the total range of subjective brightness sensations that the observer can perceive. Note that the lower (solid) portion of this subjective range is quite restricted and has a definite limit, the black point L_b . The upper (dashed) portion is not actually restricted, but, if extended too far, loses its meaning since the presence of such higher intensities would immediately tend to raise the adaptation level to a higher value than L_a .

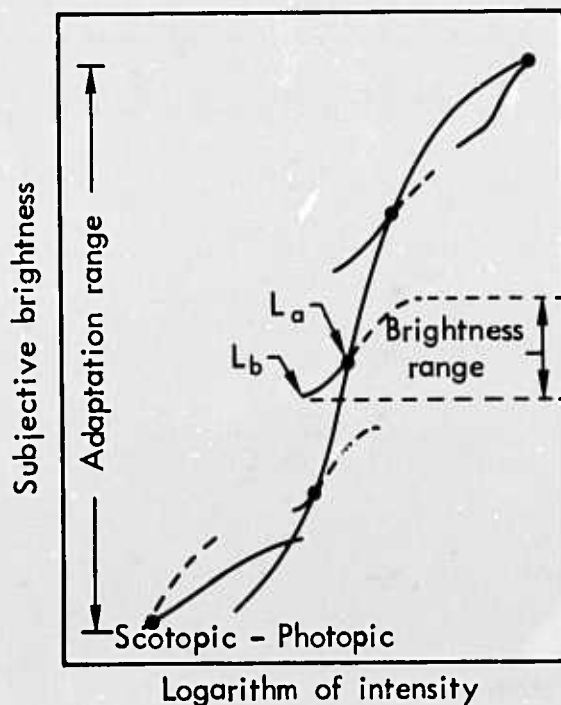


Fig.47—Range of subjective brightness sensations for
for a given adaptation level
(adapted from Ref.3)

Black, White, and Gray

In the above discussion the sensations of "black" and "white" have crept in with an oversimplified specification of the stimuli conditions for their perception. A stimulus intensity well above the black point may also appear black if it is achromatic and if there are no other achromatic stimuli of lower intensity present. If the latter are present, then the slightly higher intensity achromatic stimulus will appear gray. Now we have introduced three sensations: black, white, and gray. These sensations are just as "real," and just as much a part of human color vision, as the sensation of redness. But they are far more complex in a physical (as contrasted with a psychological or physiological) sense than any hue, and manifest less apparent connection between the stimulus and its effect than does chromaticity.

Any attempt to define a physical stimulus that will give rise to one of the sensations of white, gray, or black is exceedingly difficult. The definition is meaningful (and possibly successful) only if both the stimulus itself and the totality of surrounding stimuli are simultaneously defined. Of course, in a general way, these sensations are produced by surfaces that are spectrally nonselective with respect to the general illuminant. Under such conditions the brightest nonselective area that is not obviously a light source will tend to appear white. A light source is never perceived as white, but either as colored or colorless. The sensation of white is a positive sensation and is not the absence of color.

Any nonselective surface that is not the brightest in the field of view will tend to appear gray. Furthermore, a colored surface whose perceived brightness is lower than that of a white present in the field will appear to contain gray. Colored areas brighter than the white will not. In other words, gray is a perception that occurs when there is a brighter area in the field of view, and the same perception is produced equally well whether the stimuli are perceived as colored or not.

For example, ⁽³⁾ a spot of noncolored light that when projected on a screen appears white to the observer will immediately appear gray

if surrounded with a higher intensity area. A spot of red light, when surrounded by a brighter border of the same red, will appear to contain gray. Finally, in reference to the appearance of nonselective surfaces discussed above, a nonselective surface may be perceived simultaneously as white and as less bright than another white surface in the field of view if the latter is perceived as being more strongly illuminated. For example, the portion of a piece of writing paper in a shadow may be perceived as the same white as that portion out of the shadow.

In short,⁽³⁾ white is a phenomenon relating to the perception of a surface, gray is a perception of relative brightness sometimes of a surface and sometimes not, while black may be considered a positive perception of the inadequacy of the stimulus to produce any other visual response. The preceding is not intended to do justice to these important achromatic sensations.^{*} However, the above should serve as a vivid illustration of an extremely important aspect of human color perception: Little success can be expected for the prediction of the sensations perceived as the result of a given stimulus unless due regard is given to both the mental state of the observer and the totality of his surroundings.

Chromatic Adaptation

An equally remarkable adaptation phenomenon of common experience is that of color constancy. White surfaces and everyday colored objects tend to be perceived as essentially unchanged in color even when the illuminant differs considerably in spectral composition, as between daylight and tungsten lamps. This apparent constancy of perceived object color is reasonably successful so long as the illuminant is not strongly colored and the reflectivities of the objects viewed are not sharply limited to narrow wavelength bands. It tends to fail increasingly as these conditions are violated. It should be noted that this phenomenon is not the same as so-called memory color, although the

^{*}Evans discusses these sensations further in a recent paper on the variables of perceived color. (See Ref. 31.)

latter probably has a good deal to do with the degree of success in normal everyday color constancy.

By means of this color adaptation, the visual process tends to readjust automatically to any illuminant and object scene such that a normally white object may still be perceived as white, and the object colors in the scene may still have their approximate daylight hues. It must be stressed that for this adaptation process to occur, the scene must have detail, i.e., objects, shapes, boundaries, etc. In general, the more varied the detail, the greater the possible adaptation.

The ability of the human visual system to maintain approximate color constancy by adapting to the illuminant color is illustrated by the comparatively well-investigated daylight/tungsten case.^{(143), (144)} This example also shows the differences possible between what the CIE Standard Observer would "see" and what a normal real observer would perceive for identical stimuli. The following example is taken, with minor changes, from Burnham.⁽¹⁴⁴⁾

As noted in Section III, color is normally specified technologically with reference to the CIE Standard Observer. A fixed appearance is associated with each such specification because the specification itself implies both a constant state of adaptation and a fixed set of viewing conditions. CIE specifications, however, may not give an accurate indication of the appearance of colors under other viewing conditions. The long vector in Fig. 48 represents the large "blue-yellow" shift in CIE specification of a spectrally nonselective object when it is illuminated respectively by CIE sources C and A. A spectrally nonselective surface in everyday life, however, retains a near-neutral appearance under either daylight (source C) or tungsten (source A) illumination. The actual small change in appearance would be represented by a vector no longer than the short one in Fig. 48.

If a spectrally nonselective surface remains so nearly neutral in appearance under different illuminations, why should the CIE color specification for that surface shift so much? The reason is that the Standard Observer is assumed to look always at two adjacent areas through a

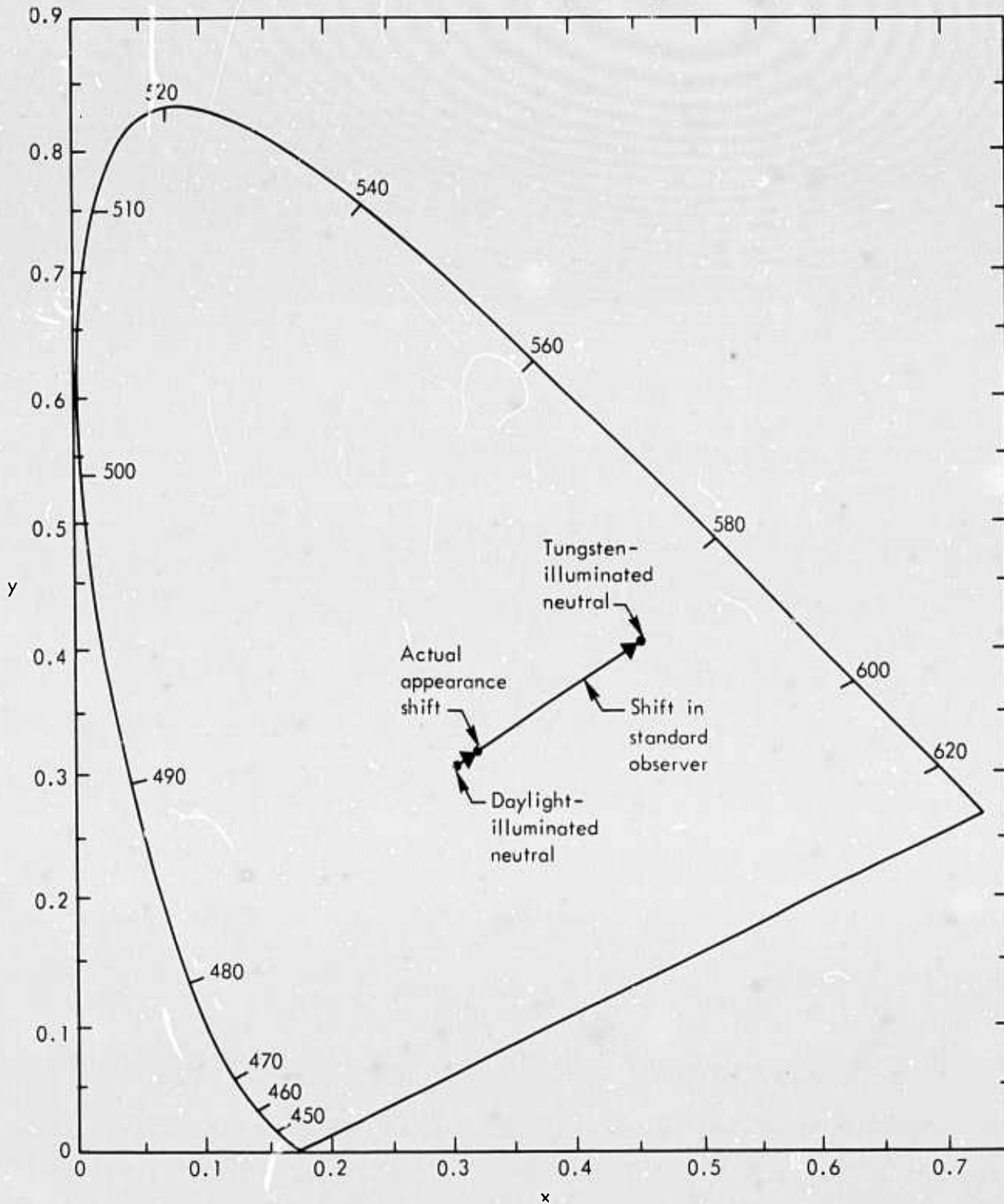


Fig.48—Comparison of actual appearance shift and shift in CIE specification for a neutral surface in daylight and in tungsten illumination (adapted from Ref. 144)

dark tube. The light illuminating either area can have no effect on the assumed constant adaptation of the (single) standard eye. Consequently, if two identical spectrally nonselective surfaces are placed side by side and one is illuminated by daylight while the other is illuminated by tungsten, they will appear about as different to the

Standard Observer as a yellow rose and the blue sky. This is about the difference in color appearance represented by the long vector in Fig. 48.

In everyday life, identical adjacent areas are not usually illuminated with widely different illuminants, nor are they looked at through dark tubes. It is more typical to find one kind of illumination incident on all objects in view. The general adaptation of the visual system is not determined by small patches in a dark surrounding area, but by the spectral quality of a prevailing illumination. When the illumination is changed from daylight to tungsten, the visual system adjusts its spectral sensitivity to such an extent that in a matter of minutes, it largely compensates for the change in quality of the general illumination. This is also true for other quality illuminants. In general, the visual system continuously adjusts its sensitivity to the prevailing general illumination, compensating for the illuminant quality in such a way that ordinary object colors, neutral or otherwise, tend to remain constant in appearance. The possible role of the middle layers of the LGN in this phenomenon was mentioned in Section VI. The ON-or-OFF cells there make no response whatever to an equal energy white stimulus and respond to a colored stimulus in a red-green, yellow-blue, opponent-color fashion. The average activity in this layer would thus reflect precisely the color quality of the illuminant for ordinary complex viewing situations and thus provide a dynamic means of "discounting the illuminant color" to obtain color constancy for perceived objects.

In addition to general color adaptation, similar phenomena occur both locally and laterally during the viewing process.⁽³⁾ By this is meant that as the point of fixation moves over a scene, the color perceived at any point is influenced both by the color perceived at the previous fixation point and by the color of the region immediately adjacent to the point of fixation. The subject of color adaptation is far more complex than the above example of general adaptation to illuminant color might portend. The observer's intention must always be considered in other than highly artificial, controlled viewing conditions. He may, on occasion, perceive the same scene in two or more

modes, each corresponding to a different state of adaptation, and he may actually switch back and forth, consciously or unconsciously, from one mode to another.

An illustration of the latter is the so-called complementary shadows demonstration.⁽³⁾ An object is placed in front of a large white screen and illuminated by two colored lights, say, yellow and red, in such a way that over the whole screen and object surface, except for the shadows, there is a mixture of red and yellow light forming an orange color. One shadow is illuminated only by yellow light, the other only by red. The sensations ordinarily perceived by an observer viewing this setup are as follows: Mode (1): When viewing the scene as a whole, the orange color of the yellow-red mixture is perceived. Mode (2): When comparing the two shadows, the yellow shadow is perceived as a blue-green and the red shadow as a purplish-red, i.e., the two shadows are of complementary hues. This phenomenon occurs for any pair of projection colors, although if the two lights differ only slightly in spectral distribution, the saturation of the complementary hues will be quite low. The phenomenon may be observed in a daylight illuminated room at an illuminant level sufficient for normal perception of other colored objects in the room. Under adaptation mode (1), above, the observer perceives the yellow-red mixture as orange, as in normal daylight adaptation. Under mode (2), he unconsciously color-adapts to the orange background, regards the orange as neutral, and perceives the two shadows as complementary hues relative to the mixture as a "white." Note that to perceive under this adaptation mode he does not have to consciously think of (i.e., perceive) the mixture color as being "white."

Local and lateral adaptation effects occur both in brightness perception and in color (hue) perception. The brightness aspect may be described as the effect that makes a surface appear brighter when it is surrounded by a large darker area than when it is surrounded by a large lighter one. In other words, the difference in the brightness of two surfaces is exaggerated if they are adjacent, the effect being a maximum when one surface is a small area completely surrounded by a large area.

The preceding might be termed a subtractive effect. Its opposite, an additive effect, is also observed. For example, if in a fairly intricate and small-scale design, the details are surrounded by narrow black lines or the design consists of narrow black lines on a uniform background, the black is apparently added to the entire design so that colors appear darker and more saturated, and whites appear grayish. If instead of the black lines, white lines are used, white is added throughout, and the colors appear lighter and less saturated. In colored illumination, if a small neutral object is placed against a large neutral background, the object appears to have the hue of the illuminant if the background is lighter than the object. Further, with non-colored light, a grey object on a colored background of the same lightness appears to have the complementary hue.

Despite the fact that the various aspects of brightness and color adaptation may cause significant changes in the visual sensations perceived, i.e., in brightness, hue, and saturation, adaptation has almost no effect on the matching of equivalent stimuli. In other words, equivalent stimuli remain equivalent regardless of adaptation changes, although the color sensation perceived as belonging to the stimuli may change drastically. Further, it is important to note that the relative luminous efficiency function V_λ does not seem to be affected by the adaptation phenomena, even after an intense color adaptation that would seem to reduce necessarily the sensitivity in the region of the adapting color.

There is, therefore, a marked difference between problems dealing with equivalent stimuli and problems concerned with brightness and color equality of sensations. Further, brightness and color adaptation seem largely independent. Thus, if one sets up a 3-receptor theory to explain the color-matching experiments and proposes changes in relative receptor sensitivity to explain color adaptation phenomena, he must then assume a fourth receptor to account for the constancy of the relative luminous efficiency function with adaptation, since the latter cannot then be a simple sum of the three receptor curves.

In point of fact, MacAdam⁽¹⁴⁵⁾ found that his extensive measurements of pairs of colors that produced identical perceptions when the

observer was adapted to two different chromaticities required the assumption of at least 5, and in one case 6, different receptors for an adequate explanation. His receptor system was assumed to be of the Von Kries type, wherein the responses of the receptors are differently affected by adaptation, but their individual relative spectral sensitivities are not.

It is interesting to note that the one case that MacAdam found to require a sixth receptor was for the greenish-yellow/daylight pair of adaptations. The greenish-yellow used had a dominant wavelength at about 560 mμ, i.e., at the maximum of the photopic luminosity curve. The additional phenomenon found with this correspondence is believed by the present writer to be more than mere coincidence. Experiments with yellow adaptation always seem to give results that differ in character from the homologous results found with all other adapting hues. Yellow adaptation seems to have little or no effect on the hue sensations perceived for given stimuli and in this sense behaves somewhat like general brightness adaptation. This achromatic behavior is consistent with the psychological fact that yellow is the least saturated hue.

THE YELLOW ANOMALY

The unique role in human color vision played by yellow, and its intimate and seemingly inexorable relationship with white (i.e., neutral or achromatic), is most vividly illustrated by such adaptation phenomena. But the unusual nature of yellow is immediately apparent with additive, trichromatic color mixtures. An excellent yellow (almost spectrally pure) is easily produced by a mixture of spectral red and green lights. The anomaly is that the mixture-yellow is "too good." No other distinct hue may be so produced with such a high purity.

The four distinct psychological hues are red, yellow, green, and blue. All other hues appear (i.e., are perceived) to be a mixture of the two adjacent distinct hues in the well-known color circle. Red, green, and blue are all psychologically distinct hues and are the best choices for a practical additive system of three primaries. But yellow,

which is producible from two of them, is psychologically distinct from any of them. It has been repeatedly stated by many authors that the human system for color vision cannot be said to be understood until the role of the ubiquitous yellow has been explained.

The rapid chromatic adaptation results of Boynton⁽⁶³⁾ offer a striking illustration of the uniqueness of yellow. Figure 49 shows his discrimination threshold curves for five different background (surround) conditions. All of the data are determined by the threshold judgment method. The background (adapting) stimulus subtended a visual angle of $4^{\circ} 44'$ and was of 0.56-second duration. The use of such short-duration adaptive stimuli should preclude effects of retinal chromatic adaptation and emphasize the rapid cerebral effects. The monochromatic test stimulus subtended $47'$ and was presented 0.05 second after onset of the background stimulus. The duration of the test stimulus was 0.04 second. All background stimuli were equated for luminance at approximately 16 mL. The data are presented in relative units, but the same units are used throughout, and hence all data in Fig. 49 are directly comparable. Each point represents the relative energy required, at that wavelength and for constant bandwidth, to achieve discrimination between the test stimulus and the surround.

Some of the features exhibited by these data are as follows:

1. All five sets of curves are virtually identical in relative sensitivity above 610 m μ . They may be made coincident in this region by a shift in ordinates. Note that this applies also to the red-adapted curve. Thus the effect of rapid chromatic adaptation is merely to change the overall sensitivity in the region above 610 m μ , regardless of the adaptive hue.
2. The red, green, and blue adaptation curves are mutually similar (homologous) in the sense that each has a pronounced maximum (the presence of secondary maxima does not obviate this similarity). But the yellow-adapted curve is quite unique in that it oscillates slightly about an approximately constant sensitivity level below 610 m μ .

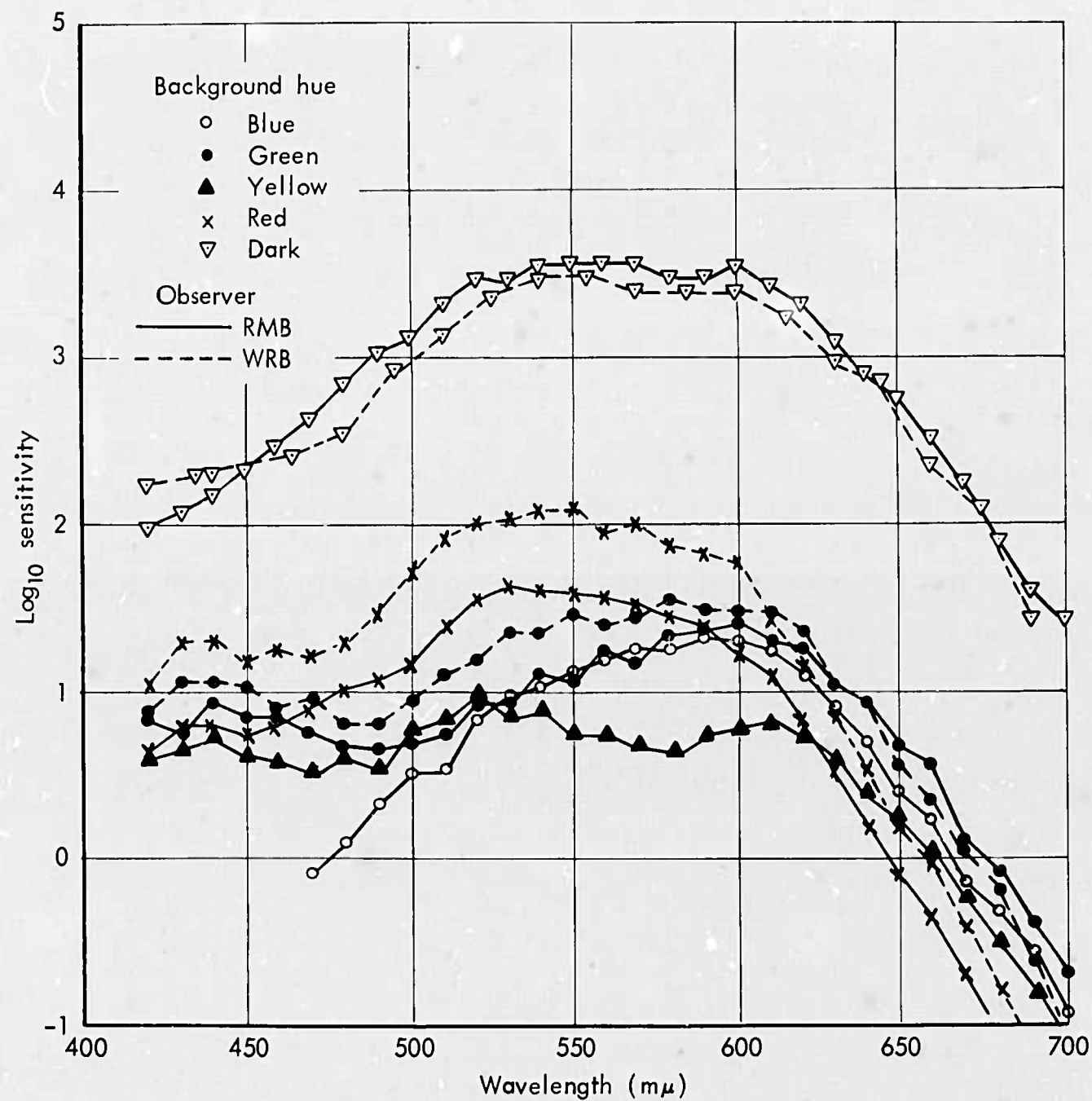


Fig.49 — Spectral discrimination thresholds under rapid chromatic adaptation to various backgrounds by Boynton (Ref.63)

3. The yellow-adapted curve has by far the lowest sensitivity level, being 0.5 log unit below the blue curve maximum sensitivity and 1.0 log unit below the red curve maximum sensitivity. The full potential selective effect for a blue background may not be indicated by Fig. 49, since the blue stimulus used had a bandwidth some five times as wide as that of the red, yellow, and green.⁽⁶³⁾

For comparison with these rapid adaptation results, it is interesting to note that Hurvich and Jameson,⁽⁶⁹⁾ using a heterochromatic brightness-match technique with completed chromatic adaptation, found for one subject no change whatever in the shape of the subject's relative luminosity function for yellow adaptation, while changes for red, green, and blue adaptation were appreciable.

The "achromatic" effect of yellow adaptation on perceived hues is shown by Boynton's results for the variation of the subjective appearance of the test stimuli with the different colored backgrounds, as indicated by subject responses. Table 1 presents a summary of the responses by one of Boynton's subjects to stimuli in the wavelength ranges noted. The unique effect of the yellow adapting surround is apparent. The yellow background case is the only one in which the complete, ordered spectrum of perceived hues was retained. The yellow background case was the only one for which no part of the spectrum was perceived as white. Finally, it is again emphasized that Boynton's results are for short-duration stimuli, and the "adapting surround" conditions of Table 1 specifically do not refer to the general chromatic adaptation discussed earlier.

Much more could be written about "anomalies" in the yellow region of the spectrum and the apparently unique relation discussed above between yellow and white. The literature abounds in both observation and conjecture on this subject. As yet, however, the data seem incapable of supporting a definitive statement either of the exact nature of the "yellow anomaly" or of its physiological origin. But it seems certain that yellow has a relationship with white not possessed by its complementary blue.

Table 1

SUBJECTIVE APPEARANCE OF MONOCHROMATIC TEST FLASHES

Adapting Surround	Wavelength (mμ)	Subject Response ^a
Red	420-460	Blue, purple
	470	White
	480	White
	490-550	White, yellow
	560-690	Yellow, white
Yellow	420-460	Blue
	470-480	Blue, green, blue-green
	500-540	Green
	550-560	Green, yellow
	570-580	Yellow
	590	Orange
	600	Orange, red, yellow
	610	Red, yellow
	620-700	Red
Green	420-480	Blue
	490-510	Blue, green, blue-green
	520-530	White, green
	540-580	White, red, green, orange, yellow
	590	Red, orange
	600-700	Red
Blue	420-520	Blue, white
	530-560	White, yellow
	570-580	White, yellow, orange
	590-660	Red, orange, yellow, white
	670-700	Red

^aResponses to a given wavelength listed in order of decreasing frequency of occurrence. Data is for one subject by Boynton (adapted from Ref. 63).

THE WHITE ANOMALY

Considerable use has already been made of the term "white," in Sections III and IV, as well as in the present section. This usage has by no means been uniformly precise. This is not entirely the fault of the present writer. In visual research, as in everyday life, the term "white" is used to ascribe diverse aspects to diverse phenomena.

Generally speaking, "white" is taken to mean a hueless sensation. This makes general sense only if we exclude white as a "color attribute of surfaces," and this we do not wish to do. Instead, we make the term dichotomous: (1) White is used to denote the positive sensation that we consciously call "white" and attribute as a characteristic of an object or a surface, as, for example, a white shirt. Such a positive white has exactly the same sensory status as does red. (2) White is also used to denote chromatic neutrality, or achromacy. In this context, it may refer to radiation whose spectral intensity is independent of wavelength, or to an object whose spectral reflectance is independent of wavelength. But more generally (and more importantly) it refers to any physical visual stimulus that an observer treats as having no hue.

The first meaning of "white" given above states that white is a sensation whose presence or absence may be simply reported by an observer. The second meaning is less clear, and it is difficult to make the meaning more precise. Therein lies the "white anomaly."

A great deal of the psychophysiological data on human vision makes sense only if the sensations of hue are considered relative to each other. The null or balance point of this relativity is the white point. It is to be emphasized that an observer need not be consciously aware of the existence of this balance point for his visual system to use it. General chromatic adaptation to a new illuminant, discussed earlier, is a case in point. So is the complementary shadow phenomena. Such phenomena are rather neatly "explained" by shifting the white point about on the CIE chromaticity diagram, as indicated in the example of Fig. 48.

The real difficulty arises when one attempts a psychophysical correlation between stimulus and sensation. Hurvich and Jameson found⁽¹⁴⁶⁻¹⁴⁸⁾ that even under the specialized conditions typical of color-matching experiments, the specification of a "white" stimulus must include the spectral distribution of the radiant energy, luminance level, stimulus area on the retina, stimulus duration, and "preadaptation" state of the observer. This represents a far more detailed specification requirement than that for, say, a "red." Furthermore, the highly transient nature of the temporal dependence of such a specification implies an enormous complexity in any dynamic description of human vision under ordinary circumstances.

The physiological mechanisms subserving this dynamic achromacy in human color vision are not known. Yet it seems safe to say that they are the most important aspects of the visual system so far as the organism is concerned. Postulated mechanisms must not be restricted to peripheral structures, as is often the case. As previously noted, the role of the LGN in setting a dynamic balance point seems highly probable. Also a definitive case for the participation of even higher cerebral processes in this phenomenon is provided in the experiments discussed below.

THE KOHLER AND LAND EXPERIMENTS

We conclude this section on psychophysiology with a discussion of some phenomena not normally associated with adaptation in its usual sense. Nevertheless, they have a definite bearing on the same underlying mechanisms. Furthermore, the experiments are particularly illustrative of the cerebral control over the visual sensations perceived by a human observer and the dynamic nature of this control.

The Kohler Experiments

Cerebral control over perceived sensations is well illustrated by the extensive experiments conducted by Kohler⁽¹⁴⁹⁾ on the psychophysiological effects produced in subjects who wore various types of optical devices over their eyes. His first experiments were with binocular inversion spectacles.

At first, two "realities" were perceived by the subject as a result of the conflict in the sensory information from his eyes, on the one hand, and from the rest of his body (muscles, joints, ear canals, etc.), on the other. In visual space, left was now right, and up was down. On the second day, the "outside world" was still upside down, but the subject's body was now felt as a standard. By the fourth day it was clear that the body's tactile receptors would win out and force the visual perceptions into agreement.

The degree of this cerebral "adaptation" and the speed with which it was accomplished are indicated by the fact that reflex reorientation was complete by the fifth day, and the subject was skiing (at Innsbruck, Austria) on the tenth day. Yet despite this perfection in "adaptation," the subject was still able, under some conditions, to perceive a duality. For example, he could perceive a weight swinging suspended on a cord as being upside down, but if he touched the supporting cord, the visual image reverted instantly to a normal pendulum movement.

These inversion experiments, while spectacular in the resulting adapting speed, actually do not involve the introduction of distortions in the visual field. The entire scene is simply inverted. In another set of experiments with several subjects, Kohler used 15° to 20° prismatic spectacles that introduced both "bends" and chromatic aberrations (rainbows at the edges of objects). The bends in the visual field were "straightened out" within ten days. If the subject then removed the device it took him four days to straighten out the aftereffects of bends in the opposite direction.

Contrasting with these relatively rapid adaptation processes for spatial effects were the results for color disturbances. The subjects had to wear the prisms for several months (up to 124 days) before they finally succeeded in compensating for the prism-induced chromatic aberrations. When ultimately the prisms were removed, chromatic deviations of opposite nature appeared as aftereffects, requiring several weeks for their removal. It is interesting to note that these aftereffects were visible even in monochromatic yellow light. The long time required for the color compensation implies the modification of a mechanism of considerable precision. (13)

These experiments vividly illustrate the human organism's ability to adapt his mental interpretations (perceptions) of the afferent nerve signals to suit his needs. His brain chooses from the available afferent signals and manipulates those so chosen to give sensations in agreement with his experience.

The Land Experiments

This ability of the human brain to make the best of what is at hand and an indication of its limitations under certain extraordinary circumstances are illustrated by the Land⁽¹⁵⁰⁾ experiments with two-primary color projections. Land prepared two color-separation positives by taking two photographs, one through a filter transmitting the long wavelength portion of the visible spectrum (585 to 700 mμ), the other through a filter transmitting the middle region (490 to 600 mμ). The exposures were adjusted to produce identical records of the gray scale (reflectance range from zero to one). The two positives, called respectively the "long-wave record" and the "short-wave record," were projected in register on a screen with various combinations of projector lights. The customary combination was red light (590 to 700 mμ) with the long-wave record and incandescent-lamp light with the short-wave record.

When the long/short records for a "natural" scene were so projected, the average viewer saw on the screen what appeared to be a projection of some sort of color transparency. In the portrait of a girl, he saw blond hair, blue eyes, a red coat, a blue-green collar, and rather natural flesh tones. For an object scene, he perceived a yellow pencil, a white pencil, a white jar, a blue-green book, a pale blue ocean on a map, a brown box, a green blotter, red printing on a magazine, etc. These color sensations were seen with no apparent time delay after presentation, and under sufficient room illumination to permit perception of ordinary colored objects in the room.

The detailed results of most of Land's experiments are not of primary interest here. Neither are his interpretations of them, which have generated considerable controversy.⁽¹⁵¹⁾ Land's work has received

careful appraisal by Judd,⁽¹⁵²⁾ who notes that the majority of Land's results are in agreement with, and are explicable in terms of, previously investigated adaptation phenomena such as those discussed earlier in this section. A similar conclusion was reached in a more qualitative fashion by Woolfson.⁽¹⁵³⁾ With regard to these well-known two-color projections, Land's major contribution has been to demonstrate the astonishing extent to which the human visual system is able to make adjustments in order to perceive object colors of all hues, even though the visual system receives incomplete color information. His experiments emphasize the extremely important role played by cerebral mechanisms in such perception.

However, two of Land's results are quite unusual. One is concerned with the failure of the two-primary color projection in a situation where success was expected (and predicted).^{(30),(152)} The other result is the emergence once again of unique phenomena in the yellow portion of the spectrum.

The failure (i.e., achromatic) cases have been discussed by Judd.⁽¹⁵²⁾ As he pointed out, Land discovered that in an object scene shown by two-primary color projections, the objects will be perceived as having essentially no hue (or no variation in hue) if the amounts, C_1 and C_2 , of the two projection primaries in all portions of the scene satisfy the relation

$$\log C_1 = a \log C_2 + b, \quad (23)$$

for any values of the constants (a , b). This achromatic condition is shown in Fig. 50, using Land's coordinate system. For comparison, the loci of subjective colors for his standard system (red projection of long record, tungsten projection of short record) are shown in Fig. 51. In the latter, the 45° line, corresponding to the locus of points in the scene for which $a = 1$ and $b = 0$ in Eq. (23), is the achromatic locus. This is as would be expected for such a scene. It represents the constant chromaticity of equal portions of the two projection primaries,

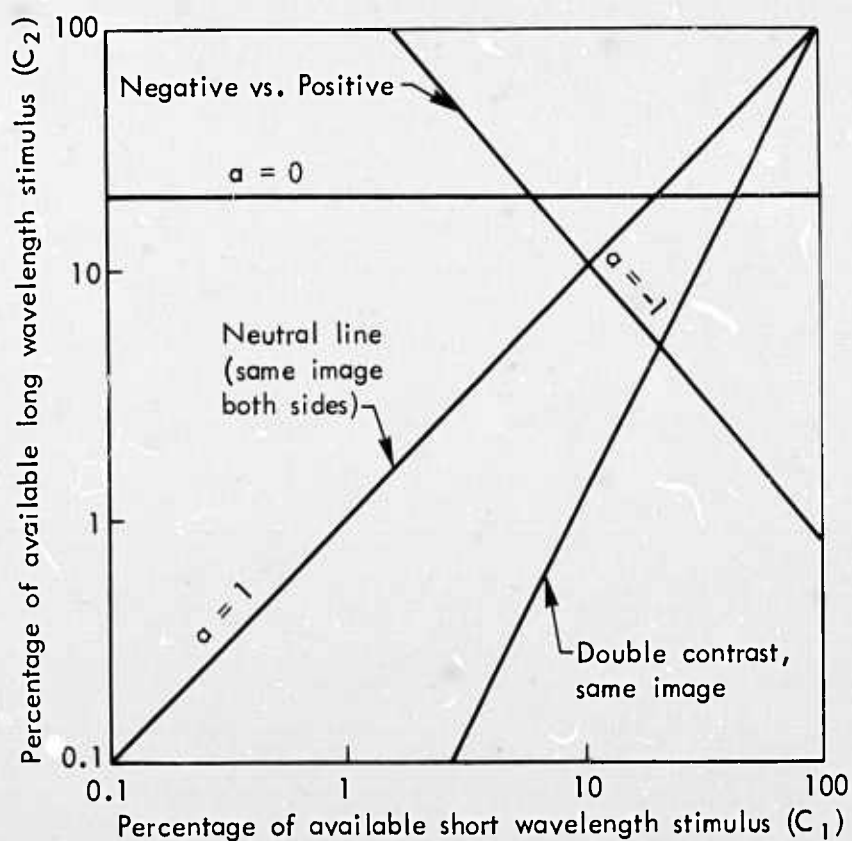


Fig.50—Loci of projection primary stimuli combinations that result in achromatic (or monochromatic) images, according to Land (Ref. 150)

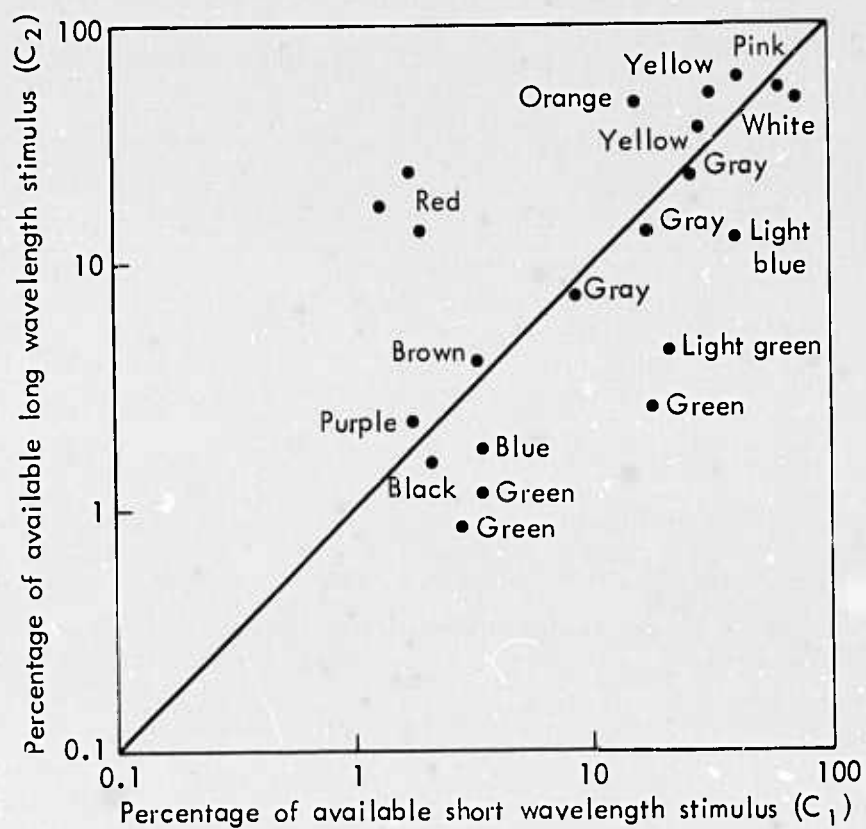


Fig.51—Loci of subjective colors for the Land standard two-primary color projection system (Ref. 150)

and hence constitutes the adapting "illuminant" for the observer in the standard⁽³⁰⁾ adaptation explanation of the appearance of the hues in Fig. 51.

On the other hand, the line $a = 1$ in Fig. 50 has a different meaning, namely, that every point in the scene satisfies Eq. (23) with $a = 1$. In this case there may still be a full range of brightness variations among the objects in the scene, but it is a variation of a single hue and is perceived as such, or as being achromatic. Judd⁽¹⁵²⁾ ascribes the failure of the two-primary system to produce color in this case ($a = 1$) to the subject's perception of this singleness of hue. A seemingly more important (and positive) result is that brightness (and hence spatial contrast) variation per se does not constitute sufficient information for the visual system to perceive a multiplicity of hues. That brightness variation is a definitely helpful factor is unquestionable, when the (not yet established) required minimum of information is available.

To depict a scene satisfying Eq. (23) with $a = 0$, one places a transparency in one projector and none in the other. The resulting scene consists of a brightness variation in a single hue for the various objects, superimposed upon a constant brightness "wash" from the second projector hue. Again the scene is perceived either as achromatic or as a brightness variation in a single hue. Judd⁽¹⁵²⁾ views this "failure" as due to the visual system's perceiving the information on the screen as coming from two distinct sources. One source is a "black-and-white" picture projected through a colored filter, and the other source is a constant brightness colored "illuminant." That is, Judd assumes that the visual system un-mixes the mixture on the screen.

But again there seems to be a positive result here. There are certainly brightness variations, as in the $a = 1$ case, over a considerable range. In addition, there are variations in the chromaticities of different portions of the screen, covering the range from one projector hue almost to the other (by varying relative intensities). In fact, as Land⁽¹⁵⁰⁾ noted, this case may be set up so that nearly the same range of primary ratios C_1/C_2 are present as in the standard multi-hue system. But this multiplicity of ratios no longer presents the

visual system with adequate information for the perception of an ordered spectrum of hues.

Now the only difference, in principle, between the $a = 0$ case and the standard "multicolor" procedure is that in the standard case, the variation in the ratio C_1/C_2 itself contains information from the original scene in an orderly form. If small values of the ratio correspond to blue-green objects in the originally photographed scene, then large values correspond to red, and the intermediate values of the ratio correspond to portions of the spectrum lying between blue-green and red. (The blue-green and red limits are determined by the choice of color-separation filters used for the original negatives.) But the brightness ratios for the case $a = 0$ do not contain any such ordering. They correspond, in fact, only to the information contained in a single color separation negative, which, standing alone, has no chromatic content.

The preceding "failures" strongly imply that Land's "successful" two-color projection techniques provide definite evidence of the visual system's utilization of the ordered information contained in the two separation negatives by virtue of the choice of the original photographic filters, Judd's neutrality⁽¹⁵²⁾ and Walls' contrary assertion⁽¹⁵¹⁾ notwithstanding.

The "failure" situation for $a = -1$ in Eq. (23) seems to afford almost conclusive proof of the existence of such a phenomenon. The $a = -1$ case results if a separation negative is placed in one projector and the positive from that same negative is placed in the other projector. With proper adjustment of the positive and negative densities and the projector intensities, a considerable range in C_1/C_2 is obtained for the "objects" depicted in the scene. The formulation⁽³⁰⁾ used by Judd⁽¹⁵²⁾ to explain the multiplicity of hues in the standard Land system predicts a similar color variety here. But none was found.⁽¹⁵⁰⁾ If red and tungsten are the projection colors, the observer perceives only a red of various saturations through pink to white.

Neither Judd⁽¹⁵²⁾ nor Walls⁽¹⁵¹⁾ offers an explanation of the $a = -1$ case. Judd suggests that the explanation might lie in the observer's perception of regularity (i.e., one-dimensional variation)

among the colors of the light patches making up the scene, instead of the object color chromaticity otherwise expected. This is petitio principii, using Eq. (23) to explain itself.

As Judd himself notes,⁽¹⁵²⁾ his "explanations" of the $a = 1$, $a = 0$, and $a = -1$ failures are all different. He also states that Land's experimental result, embodied in Eq. (23), is both new and deserving of serious attention. The present writer feels that the failure of the visual system for cases satisfying Eq. (23) is completely explained by the loss of the ordered information encoded via the original photographic filters. It is precisely this information that is deleted in the reduction from Land's standard two-dimensional system to any one-dimensional system represented by Eq. (23). The nature of the "ordered information," the photographic "encoding" procedure, and possible physiological mechanisms subserving the utilization of the encoded information are discussed below. But it seems reasonable at this point to conclude that the experimental evidence indicates a new phenomenon that we shall call the "Land Effect," defined as the ability of the human visual system to utilize a specific two-dimensional code that correlates perceived variations in the projected scene with object hue variations in the original scene.

The Land Effect is illustrated in its purest form by Land's so-called yellow-yellow experiment.⁽¹⁵⁰⁾ With a dual monochromator projection system (projecting the short-wave record at 580 m μ and the long-wave record at 590 m μ), a surprising range of perceived hues is obtained (Land states, "the full gamut," as indicated in his color map, Fig. 52). Under such conditions the contribution of normal color adaptive mechanisms is surely minimized (though still present), and the contribution of the specific information encoded in the ratio C_1/C_2 is maximized. This case also provides the first direct evidence about the encoding procedure that must be used in taking the original color-separation negatives, in order for the information so recorded to be properly utilized by the human visual system.

First, note that for the visual system to be able to use either its normal adaptive mechanisms or any C_1/C_2 ratio information encoded in the projection transparencies, it must be able to clearly discriminate

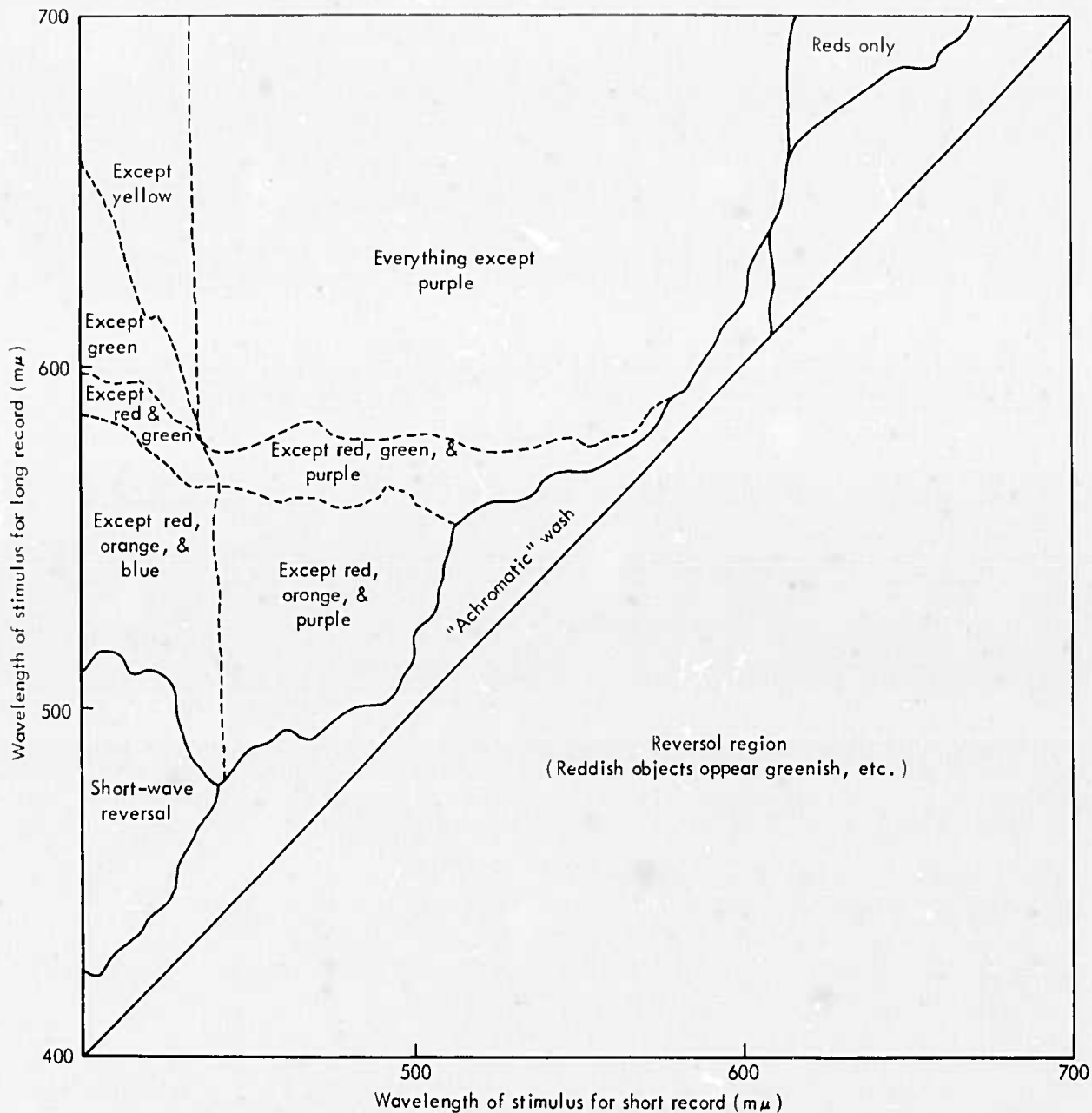


Fig.52—Land's color map indicating the range of colors perceived with different pairs of projection wavelengths in his two-primary system (Ref. 150)

the hue pair used in the two projectors. Insofar as the spectrum is concerned, human color discrimination is by far the most sensitive in the yellow, from about 570 to 580 mμ. This has been found by many investigators. The recent data of Siegel and Dimmick,⁽¹⁵⁴⁾ presented in Fig. 53, show the increased sensitivity in the yellow very strikingly. Ample discrimination sensitivity is available for Land's yellow-yellow experiment.

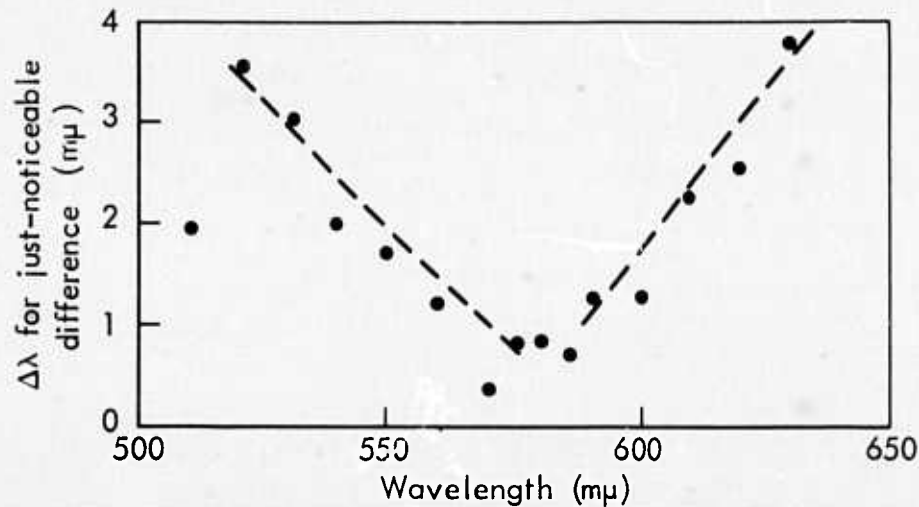


Fig.53—Color discrimination sensitivity as a function of spectral wavelength (adapted from Ref.154)

Now let us see what information is encoded in the two color-separation negatives. The spectral transmittances of Land's separation filters are shown in Fig. 54, and their product with the spectral sensitivity of Kodak Panatomic-X film in Fig. 55. If we neglect for the moment the yellowish-green to orange (550 to 600 mμ) region of the spectrum in Fig. 55, we note that the response of the long-wave record decreases roughly linearly with increasing wavelength into the red, and the response of the short-wave record decreases roughly linearly with decreasing wavelength into the blue. Hence, for objects in the original photographed scene that have spectral reflectances with pronounced maxima in, say, the long-wave end of the spectrum, there is encoded in the long-wave record a correlation between decreasing intensity of the long-wave projection hue and increasing dominant wavelength for objects in the original scene. Similarly, in the short-wave end, there is encoded in the short-wave record a correlation between decreasing intensity of the short-wave projection hue and decreasing dominant wavelength

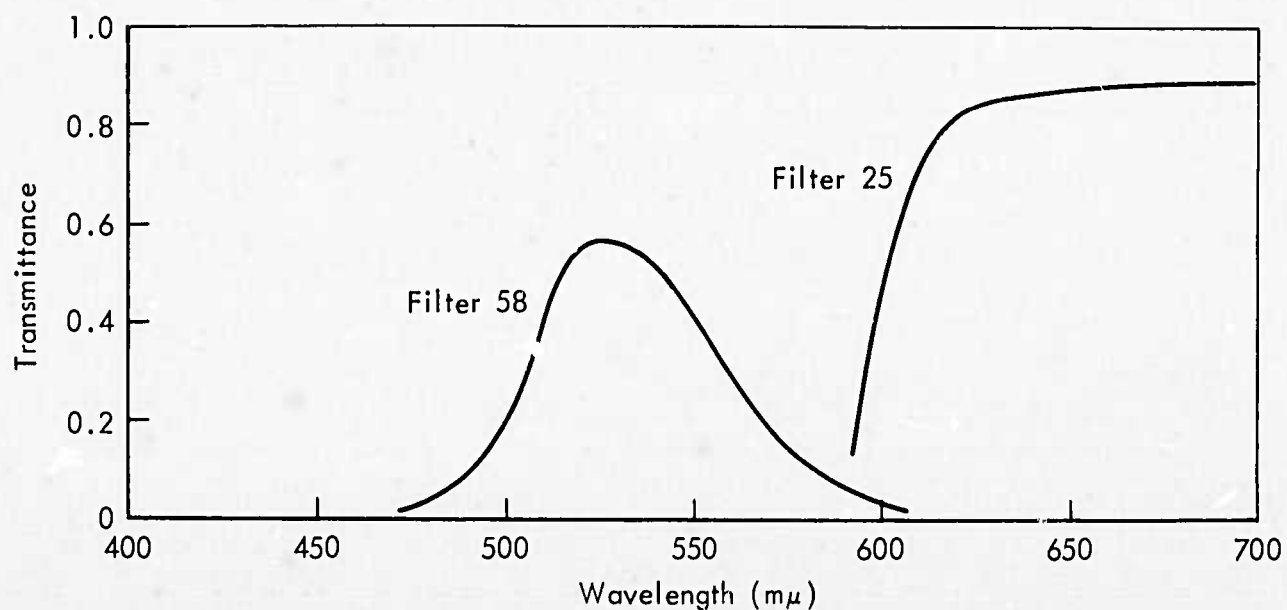


Fig.54—Spectral transmittances of Land's two color-separation filters used in making the long-wave record (filter 25) and the short-wave record (filter 58)

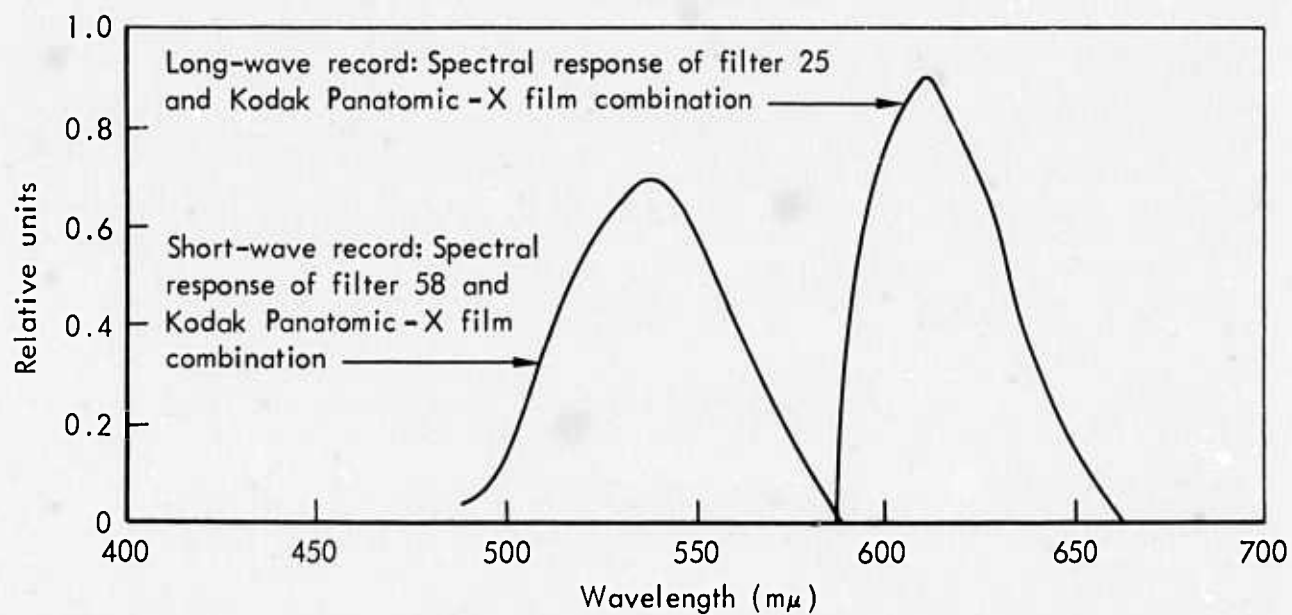


Fig.55—Spectral responses of Land's long-wave and short-wave records

for objects in the original scene. These correlations should be well isolated in the two records for these opposite ends of the spectrum, since there is good separation of the two record responses, as shown in Fig. 55.

For the mid-region of the spectrum there is little likelihood of any general correlation for photographs of ordinary scenes, since most objects, natural or artificial, have fairly broad spectral reflectance curves and will affect both records about the same. They would thus tend to produce points on or near the achromatic locus in the final projection scene, regardless of their color in the original scene. For objects with narrow peaked reflectances in the mid-region of the spectrum there will be a correlation: In the projected scene, a decrease in intensity of the short-wave projector hue, coupled with a simultaneous increase in intensity of the long-wave projector hue, corresponds to increasing wavelength in the original scene. This is precisely the same type of hue variation produced by normal adaptive effects, and hence the two phenomena may correlate to give an enhanced resultant.

Land's plot of subjective sensations in Fig. 51 seems to give definite evidence of the visual system's utilization of the above correlations. The ends of the spectrum (red and blue-green) are reached by a decrease in the intensity of the long and short projection hue, respectively. The bright mid-region follows the normal type of hue variation associated with adaptive effects.

In Land's yellow-yellow experiments, where normal adaptive effects are minimal, the visual system can still detect the wavelength correlation in the brightness variations on the screen. The variation in chromaticity for all mixtures of the two yellow projection hues is small indeed and lies on the spectrum locus in the CIE diagram. But the two projection hues may be discriminated, as evidenced by Fig. 53. It seems quite reasonable that in this case the normal adaptive mechanisms are passive. The human visual system perceives directly the correlations between decreasing intensity in the short-wave record and decreasing object dominant wavelength, and between decreasing intensity in the long-wave record and increasing object dominant wavelength. It

uses these perceived correlates to effect sensations spanning the spectrum of hues.

From Fig. 55, this intensity-wavelength correlation in Land's yellow-yellow experiment may be sketched as shown in Fig. 56.

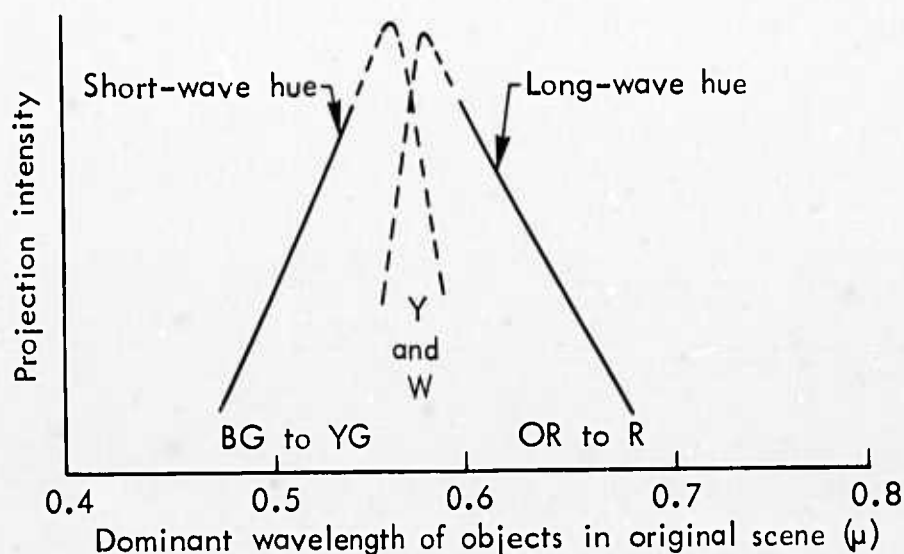


Fig.56—Proposed intensity-wavelength correlation in Land's yellow-yellow experiment

Despite the sharp dip in the yellow region of Land's photographic filter/film combination shown in Fig. 55, there is some overlap in long and short hues in Fig. 56 by virtue of Land's printing technique. He noted that the transparencies printed from the original separation negatives were deliberately made "thin." Such an overlap would enhance the ability of the visual system to use the "code" in Fig. 56, since it would provide an achromatic region in the scene. It is also interesting that Land's subjective "fulcrum" between 570 and 580 $m\mu$ is located at the pronounced dip in Fig. 55.

It is important to note that the encoding procedure discussed above is not that suggested by Land⁽¹⁵⁰⁾ in his interpretation of his results. Land felt that it was simply necessary to have an unambiguous ratio of C_1/C_2 assignable to each region of the spectrum, e.g., a monotonically increasing function of wavelength. The encoding procedure

proposed herein views such a variation as the worst choice, since this would in effect eliminate one-half of the intensity-wavelength correlation shown in Fig. 56. This would nullify the encoding procedure and hence reduce the projected scene to an equivalent one-dimensional variation analogous to the cases satisfying Eq. (23). One would expect, under these conditions, to get only mixtures of the projection hues without enhancement by the Land Effect. This is precisely what Campbell⁽¹⁵⁵⁾ found when he attempted an improvement over Land's standard procedure. If it were true that an unambiguous ratio of C_1/C_2 is responsible for the Land Effect, then, Campbell argued, the color-separation filters shown in Fig. 57 should be a definite improvement over Land's choice, shown in Fig. 54. As can be seen from Fig. 57, the filter combination 38A/106 gives a good approximation to a monotonically changing ratio of C_1/C_2 , with wavelength over nearly the entire visible spectrum. Actually Campbell found⁽¹⁵⁵⁾ that this choice gave poorer results than any other method tested.

Thus it appears that an intensity-wavelength correlation like that shown in Fig. 56 is necessary to evoke the Land Effect as herein defined. The ability of the human visual system to utilize such a correlation is not altogether surprising. The near-Gaussian shape of the human photopic luminosity curve is obvious, and the human brain is presumably well attuned to the fact that equal energy colored stimuli spaced across the visible spectrum elicit a near-Gaussian distribution in brightness sensations. Now the photographic encoding procedure used in the successful Land experiments is also essentially Gaussian, as Fig. 56 shows. Thus the conclusion is that the human visual system can (and will) utilize this particular "code" to ascribe appropriate hue sensations when viewing the projected scene in the Land type experiments. The hue information from the original scene is photographically encoded in the long-wave and short-wave records in such a way that the two records jointly have a "built-in" Gaussian luminosity function. The cerebral mechanisms "recognize" this familiar code and employ it to produce sensations in accord with experience. The only required chromatic content in the projected scene is that the two projection

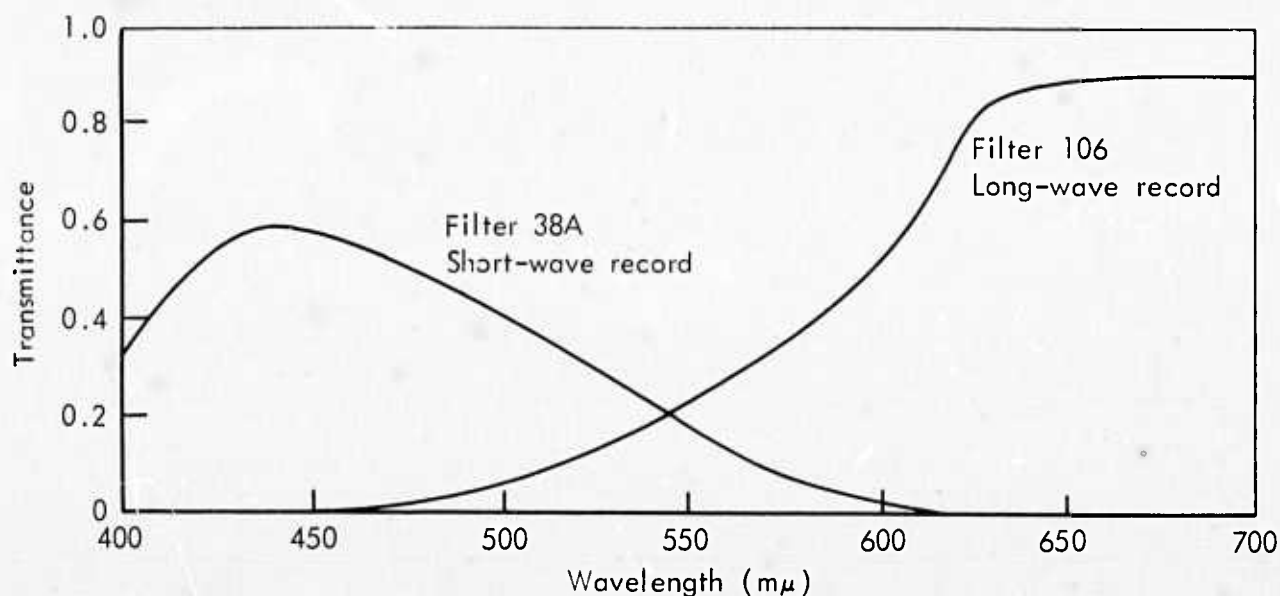


Fig.57—Spectral transmittances of Campbell's two color-separation filters

hues be discriminable so that the two "halves" of the visible spectrum are differentiated. This requirement was satisfied in all the successful Land experiments, including the yellow-yellow experiment, as previously noted.

The present explanation implies that the color space that the Land Effect can evoke is strictly one-dimensional: It can produce only a "spectrum" of hue sensations. It cannot produce variations in brightness, since it is directly dependent on such variations to produce the hue variation. It cannot produce variations in saturation, since information on saturation is not (and cannot be) explicitly encoded by means of the photographic procedure used. Finally, the Land Effect cannot produce the nonspectral hues (i.e., the purples), since any such hues in the original scene will "encode" into the mid-region of Fig. 56 and thus evoke an achromatic sensation. This conclusion is supported by Land's "sensation loci" in Fig. 51, which shows "purple"

essentially on the achromatic axis, and by his "color map" in Fig. 52, which indicates considerable difficulty in producing purples.

When interpreting Fig. 51 in the context of the preceding discussion, one must recall that it is a summary plot of many experiments, some of which involved more than the Land Effect. The latter is present by itself only if the two projection hues are not appreciably different. In Land's "standard method," which employed a red filter with the long-wave record and unfiltered tungsten light with the short-wave record, this condition is not satisfied; and other visual mechanisms are active (e.g., color adaptation, color contrast, etc.), which may couple with the Land Effect to enlarge the resultant color sensation space produced. As indicated in Fig. 52, the single-projection-hue criterion for the pure Land Effect is best satisfied for a concomitant "spectrum" of hue sensations when the projection hues are in the yellow as in Land's yellow-yellow experiment.

The present explanation of the Land Effect obviously accounts for the simple color reversal obtained by Land when the long-wave and short-wave records are interchanged in the two projectors. The "luminosity code" in the records is essentially symmetric, and the visual system accepts the record projected with the longer wavelength as the long-wave record. The "short-wave-reversal" region shown in Fig. 52 is more complex, since it nowhere satisfies the criterion for a pure Land Effect. However, it can be seen that the short-wave projection hue for this region is always well in the violet (i.e., bluish purple),* which psychophysically is a mixture of red and blue. Since all stimuli in this reversal region are confined to the short wavelength end of the visible spectrum, the visual system would favor the "red" content of the short-wave projection hue as implying the long-wave record in any successful attempt to utilize the Land Effect.

The Land Effect presents several implications with respect to the physiological mechanisms that subserve it. First, the very existence of the Land Effect implies that the suprarretinal signal complex normally contains much more nonspatial, nontemporal information than is

*Compare Fig. 13.

contained in the three-vector of the trichromatic theory. That such information is transmitted at least to the level of the LGN was noted in Section VI. Second, the Land Effect implies that the central mechanisms can "unconsciously recognize" the external luminosity function with its neutral (i.e., white or achromatic) midpoint and use it to "decode" the information in the projected scene. This is surely a reasonable extension of the activity of the central mechanisms from their analogous roles in using "unconscious neutral points" in the phenomena of complementary shadows and illuminant adaptation, discussed earlier. Third, the essentially instantaneous presence⁽¹⁵⁶⁾ of the Land Effect in a projected scene implies that relatively slow phenomena, such as photopigment bleaching, are not involved. In fact, it is important to note that the question of cone photopigments does not bear directly on the mechanism proposed above. The pure Land Effect involves essentially single-hue stimuli. No cone photopigments are likely to have extremely narrow spectral sensitivity curves. And the presence of photopigments with broad sensitivity curves would not affect the proposed mechanism even if they constituted the subsidiary mechanism permitting the required discrimination of the two projection hues. Such a mechanism could equally well be provided by a variation in cone outer segment diameters, as noted in Section V.

Finally, it is emphasized that the present treatment by no means implies that the Land Effect is primarily responsible for the perceived color variations in normal two-color projection. On the contrary, a significant part of the variation would here be due to the normal adaptive properties of the visual system. The Land Effect contributes only if the original long/short records are properly encoded.

SUMMARY

The discussion in the present section is hardly adequate to indicate the extent and importance of adaptation phenomena in human color vision. In Hering's words, as quoted by Helson⁽¹⁵⁷⁾: "Approximate constancy of the color of objects in spite of large qualitative and quantitative changes in general illumination of the visual field is

one of the most striking and important facts in the domain of physiological optics. Without this approximate constancy a piece of chalk on an overcast day would appear as dark as a piece of coal on a sunny day."

Adaptation is not something subsidiary to the primary visual system, to be explained in an after-the-fact fashion when the postulates for the "principal" mechanisms are completed. Adaptation is the feature that makes the visual system workable in everyday life and hence useful to the organism. A feature of this importance is unlikely to reside principally in peripheral receptors. The relatively slow changes involved in the illuminant adaptation phenomenon and the brightness-match color-adaptation tests may indeed involve retinal changes of some importance. But a complete explanation of even these phenomena does not lie there.

The data of Boynton, on the other hand, present large-magnitude adaptive effects far too fast to be accounted for by changes in receptor sensitivities. Helson⁽¹⁵⁷⁾ notes that adaptive effects have been found in exposures as short as 1.7 milliseconds, indicating the presence of almost instantaneous adaptation processes. The Land Effect is still present in exposures as short as 1 microsecond.⁽¹⁵⁶⁾ Under such conditions, "adaptation" must reside entirely in the processing of the afferent signal complex.

Contrast effects (also briefly mentioned above) must be viewed as dependent on the establishment of gradients with respect to the adaptation level of the visual system and not with respect to actual physical features of the contiguous objects in the field of view.⁽¹⁵⁷⁾ Ditchburn (as quoted in Ref. 158) has suggested that the colors we see may be determined very largely by edge effects occurring at the boundaries separating one area of color from another in the retinal image. The interior regions would be "filled in" by the brain. Such an attitude looks for a unified explanation for phenomena often attributed to diverse and largely independent mechanisms.

Perhaps the best way to summarize the potpourri in this section is to emphasize three facts that are recurrent in the seemingly diverse phenomena:

1. In its dynamic aspects, the human visual system displays a continuous effort to reset its various thresholds to a pattern whereby it is most sensitive to deviations from a current "equilibrium" level.
2. Insofar as color vision is concerned, the stimulus complex treated as achromatic is the principal determinant for this equilibrium level. The actual spectral distribution of this "white" stimulus is important, not just its CIE chromaticity.
3. In a dynamic utilization of the preceding two features, the central nervous system always attempts to perceive order in the afferent signals consistent with past experience. Upon finding even a minimal content satisfying such an order, it will attempt to reject the disordered content and accept and enhance the ordered content to result in familiar sensations.

The preceding statements are entirely consistent with the fundamental postulates of general adaptation level theory⁽¹⁵⁹⁾ governing the behavior of the organism. The embodiment of the general concepts of behavioral adaptation in human vision is perhaps put most succinctly by Dr. Adelbert Ames, as quoted by Judd⁽¹⁵²⁾: "What the eye sees is the mind's best guess as to what is out front."

VIII. CONCLUSIONS AND RECOMMENDATIONS

In this final section the principal conclusions reached in this study will be summarized and their implications for further research on human color vision will be discussed.

SUMMARY OF PRINCIPAL CONCLUSIONS

In Section II, "The Process of Color Perception," a working definition of the overall experimental problem was advanced. It was noted that there is a rather sharp division of effort within the science of color vision: (1) research aimed at a fundamental understanding of the physiological processes mediating color vision in humans; and (2) research aimed at the improvement of techniques for an objectively defined and reproducible system of color specification in technology. The latter may be carried on without regard to progress in the former, and without specific regard for the color sensations perceived by any given individual.

In Section III, "The Standard Observer," the terminology and experimental origin of colorimetry were discussed as embodied in the CIE Standard Observer. Particular emphasis was placed on the distinction between the experimental facts established in the matching experiments and the formalism of colorimetry derived in part from these facts. It was stressed that the results of the matching experiments specifically state that it is impossible to match the sensations produced by all radiant stimuli with additive mixtures of fixed, real, primary light sources. The colorimetric specification in such terms is accomplished only by means of imaginary primaries.

In Section IV, "Retinal Sensitivity," the essential duality of the human retinal receptor mechanism was discussed. Scotopic threshold sensitivity, mediated principally by the single rod mechanism, was found to be smooth and in reasonably good agreement among individuals. A representative portion of the available experimental data on individual foveal spectral sensitivity was collected and presented in graphical form. From a careful consideration of these data and the metameric results of Section III, the following conclusions were drawn:

1. Irregularities in V_λ curves may be anatomically based in individual subjects, but their exact origin is not established. A given irregularity at a specified spectral location may be present in some subjects and absent in others. When present, its appearance in a V_λ determination is enhanced by the use of monochromatic stimuli and small test fields.
2. The nature of small-field dichromacy results indicated their connection with the absence of a sufficient number or variety of cone receptors in small fields and suggested an integrative and/or comparison role at the bipolar level. A group of contiguous cone receptors are required for normal hue discrimination, and there is a lower limit in size for this group to function properly.
3. All attempts to fit luminosity data with three or four fundamental cone mechanisms, each with a fixed spectral response curve, have failed. The experimental results indicate that the brightness response mechanism is neither completely separate from the chromatic mechanism nor a simple summation of three or four chromatic response functions.
4. Linearity in luminance and metamerism, as embodied in Abney's Law and Grassman's Laws, is the exception rather than the rule insofar as the human visual mechanism is concerned.

Section V, "Retinal Physiology," represented an even greater distinction between the present study and most previous interpretative reviews, both with respect to the material included and the conclusions reached. Experimental data were discussed concerning the initial photoreception and the anatomy, histology, morphology, ontogeny, and electrophysiology of the retinal neurons. The established role of rhodopsin in the rods as mediator of human scotopic vision was presented as a datum for comparison of the data on cone receptors. Continued attention was drawn to the multiplicity of results indicating a fundamental difference between receptor mechanisms in the rods and cones. The following conclusions were reached:

1. No unequivocal evidence is available demonstrating the existence of three or four specific cone photopigments for the mediation of hue discrimination in human vision. The approximate trichromacy of metamerism is still the principal evidence supporting the widespread belief in a three-pigment mechanism.
2. The demonstrated existence of energy propagation by means of waveguide modes in human cone receptors presents these receptors with a spectral discrimination mechanism that must be considered irrespective of their photopigment content.
3. The retinal cone mechanism is not afferently convergent in its neural pathways, as is widely assumed. In the central retina, where color vision is fully developed, each cone has three bipolars directly associated with it. The reduction of this bipolar/cone ratio below a value of 3:1 in the retinal periphery is well correlated with the zones of reduced color vision established by psychophysiological measurements. This bipolar tripartition portends a more complex involvement of the bipolar cells in hue discrimination than present theories assume.
4. Evidence from lower vertebrates indicates that the rod and cone receptors have quite different metabolic organizations and photochemical systems and that the electrical responses at the bipolar level are graded DC potentials rather than spike discharges.
5. The cumulative implication of the above conclusions is that the peripheral mechanism subserving human color vision is not based simply on three cone photopigments, but rather is based on a combination of variation in cone physical characteristics (including any pigments) and multiple bipolar analyses. This implies that the Young-Helmholtz three-pigment theory, the Hering opponent-colors theory, and the various stage or zone theories may each contain some element of truth, but that none is individually capable of providing a suitable explanation.

The comparatively short review in Section VI, "Cerebral Physiology," on central neural mechanisms is indicative of the lack of data in this area and not the lack of their importance in color vision.

However, two important conclusions were reached:

1. The three layers in the primate lateral geniculate nuclei appear to subserve synaptically three types of ganglionic activity. The dorsal layer exhibits narrow-spectral-sensitivity, multireceptor type responses; the middle layer exhibits opponent-colors type responses; the ventral layer displays luminosity type responses. A complex, dynamic role of the LGN in human color vision seems definite.
2. The terminal synapses for the visual process are in the striate area of cerebral cortex. However, the signal complex there is not afferent determinate, but under the dynamic control of still higher brain centers. Further, the presence of a signal complex in the striate area is not synonymous with the perception of a sensation.

Section VII, "Psychophysiology," was introduced with a short review of the enormous literature on adaptation phenomena pertinent to color vision. More specific attention was then given to the unusual roles of yellow and white in both adaptation and color-mixture data. Next it was noted that there is overwhelming evidence that adaptive data generally require cerebral rather than retinal explanation. The possible significance of both the Kohler and the Land experiments relative to the cerebral "interpretations" of afferent signals was discussed. This led to a proposed definition of the "Land Effect" in terms of the "information" content of his long/short records. Finally, the rather diverse psychophysiological phenomena were summarized into three generalized conclusions:

1. In its dynamic aspects, the human visual system displays a continuous effort to reset its various thresholds to a pattern whereby it is most sensitive to deviations from a current "equilibrium" level.

2. Insofar as color vision is concerned, the stimulus considered as white (or achromatic) is the principal determinant for this equilibrium level. The actual spectral distribution of the "white" stimulus is important, not just its CIE chromaticity.
3. In a dynamic utilization of the preceding two features, the central nervous system always attempts to perceive order in the afferent signals consistent with past experience. On finding even a minimal content satisfying such an order, it will attempt to reject the disordered content and accept and enhance the ordered content, to enable the perception of familiar sensations.

DISCUSSION AND RECOMMENDATIONS

The general conclusion reached herein is that the available experimental evidence does not clearly dictate the fundamental physiological processes mediating human color vision.

The data do clearly indicate that there has been an undue concentration on attempts to "explain" the whole of color discrimination by means of a triple mechanism resident in the cone receptors themselves. The continued lack of unequivocal evidence supporting such a postulate, together with the demonstrated existence of multiple mechanisms at both the bipolar and lateral geniculate nuclei levels, portends a shaky future for simple 3-receptor theories. It must be remembered that we are still uncertain as to how much information is encoded in the ganglion cell responses. Finally, there is no definite evidence that the approximate trichromacy of metamerism must be wholly retinal in origin, as is generally assumed.

Insofar as future research is concerned, there seem to be at least three distinct fields of endeavor within which advancements are not necessarily coupled. The first is colorimetry, where the goal is an ever more adequate system for the specification of visual stimuli for use in color technology. To retain a practical system, progress toward this goal must increasingly dissociate itself from the visual sensations perceived by individuals. Advancement in this field will likely be evidenced by refinements in the CIE System.

The second field is visual biophysics, where the goal is a fundamental understanding of the physical and physiological processes that actually subserve the human visual system from reception to perception. Here, as Brown and Wald aptly remarked⁽⁹²⁾: "The problem is not to invent further models that might account for color vision, but to find the actual mechanisms that serve this function." The conclusion of the present study is that progress toward this goal is still very much in its infancy. In LeGrand's words⁽¹⁵⁸⁾: "The one certainty that emerges at the moment is the complexity of the visual processes and their interactions at all stages." Advancement in this second field would be materially aided by less concern with attempts to "prove" hallowed concepts and more concern with the establishment of reliable experimental fact. It is the author's belief that perhaps the surest road to success in this field lies in a concentrated program involving a group of subjects from birth to death. A life-spanning program that included the myriad of psychophysical tests coupled with a detailed post-mortem histology seems to be the only unambiguous way to account for intersubject variability.

There is yet a third field of human color vision research that is related to both of the preceding, but still can be pursued with considerable independence from them. This field can be called the "psychophysics of color vision" in the general sense of the term. It is concerned with the quantitative prediction of the actual color sensations perceived by a given individual under specified conditions. This implies the creation of a model whose function is to prescribe a causal relationship between input stimuli and resultant sensations. It is obvious from the discussion in this study that such a model must be far more complex than the formalism that is satisfactory for colorimetry. At the same time, it need have no pretense of representing actual physiological processes. In this connection the analogy between continuum fluid dynamics and molecular physics is apt.

The major limitation on advancement in this third field is the absence of widely accepted "axioms" upon which to build a model of the visual system. Two often used axioms are the assumption of three (or

four) specific photopigments in the cones and the law of specific energies for all retinal neurons. The lack of evidence supporting the former assumption and the considerable experimental evidence against the latter are the likely reasons for their failure to produce more satisfactory results.

The present work does indicate several postulates that may lead to fruitful psychophysical models:

1. The specification of stimuli solely in terms of CIE coordinates must be abandoned. At least some knowledge of the spectral distribution of a stimulus will be required in the model.
2. The ability of a stimulus to produce a color sensation should be measured in terms of its deviation from some equilibrium stimulus, and the actual sensation should be causally related to the details of the deviation.
3. The equilibrium stimulus should be related to the stimulus currently perceived as white or achromatic, but allowance must be made for simultaneous "effective" equilibrium stimuli, all of which need not be consciously perceived by the subject.
4. A successful model must be able to predict more than one sensation for a given input (i.e., a different sensation for each of the currently active "effective equilibrium stimuli"), each prediction thus representing a possible perception, depending on the observer's intention.

In setting up a model based on these general postulates, the established experimental evidence should be used as a guide. Certainly, mechanisms contrary to the available evidence should be avoided. In particular, hypotheses requiring an elaborate internal structure in the cone receptors are neither required nor desirable. In Talbot's words⁽¹⁸⁾: "Linkages rather than substances more simply express the mechanistic aspects of color vision." The anatomical riches available at the bipolar and lateral geniculate levels, together with the variation in physical characteristics of the cone receptors, should be the principal guides in postulating such linkages.

APPENDIX

RADIANT ENERGY TERMINOLOGY

The following defined quantities relating to the propagation of electromagnetic energy are stated without reference to any particular portion of the spectrum. They are total radiation quantities in the sense that they refer to the total electromagnetic energy contained within some finite wavelength band. The same definitions are used to denote spectral radiation quantities. That is, values of the quantities for a particular wavelength λ can be obtained by affixing the subscript λ to each quantity that contains an energy unit and modifying the units of the quantity by (unit wavelength interval)⁻¹. Thus the spectral quantities refer to a small but still finite wavelength interval containing the characterizing wavelength λ . The total quantities for an extended wavelength interval are thus related to the spectral quantities by

$$U = \int_{\lambda_1}^{\lambda_2} U_{\lambda} d\lambda, \quad (A-1)$$

where $d\lambda$ is measured in terms of the unit wavelength interval chosen for the definition of U_{λ} . The symbol U refers to the quantity of radiation, i.e., to the total energy, in ergs, contained in the interval (λ_1, λ_2) .

Radiant Flux P: Rate of energy U being emitted, transferred, or received in the form of radiation,

$$P = \frac{dU}{dt}, \quad (A-2)$$

measured in power units, ergs/second, or watts.

Radiant Intensity J: Refers to a source of radiant energy and is the radiant flux emitted per unit solid angle in a given direction

θ--

$$J_{\theta} = \frac{dP_{\theta}}{d\omega}, \quad (A-3)$$

measured in ergs/second-steradian (sterad), or watts/sterad.

Radiance N: Refers to a source and is the radiant intensity per unit projected area of the source in the given direction θ --

$$N_{\theta} = \frac{dJ_{\theta}}{dA_{\theta}}, \quad (A-4)$$

measured in ergs/second-sterad-cm², or watts/sterad-cm².

Irradiance H: Refers to a surface receiving radiant energy, and is the incident radiant flux per unit area of surface--

$$H = \frac{dP_i}{dA}, \quad (A-5)$$

measured in ergs/second-cm², or watts/cm².

LUMINOUS ENERGY TERMINOLOGY

Light is that aspect of electromagnetic energy that elicits a visual sensation when incident on the retina of the eye. Only electromagnetic energy in a narrow band of wavelengths, normally from about 380 to 750 mμ, is capable of producing such an effect, and the abilities of equal amounts of energy in different portions of this interval to produce a sensation are not equal. The reaction of an idealized average eye possessing a spectral sensitivity dependent only on wavelength as described by its relative luminous efficiency function, a numeric, is used to relate radiant and luminous quantities. The photopic quantities are obtained by using the CIE 1924 Photopic Luminous Efficiency Function V_{λ} . Thus, for each spectral radiant quantity above, there is a corresponding spectral luminous quantity, the latter being given by the product of the former and V_{λ} . The total luminous quantities are found by integration as before, since V_{λ} is dimensionless.

To indicate that the function V_λ is involved (since the electromagnetic energy is obviously unchanged), a different "energy" unit for the luminous quantities has been chosen. In relating luminous to radiant quantities, the function KV_λ is used, where K is the absolute luminous efficiency; $K = 678.8$ lumens/watts, based on the CIE defined new candle, or candela. In the following, the spectral quantities have the factor (unit wavelength interval)⁻¹ in their units as before.

Luminous Flux F:

$$F_\lambda = KV_\lambda P_\lambda, \quad (A-6)$$

$$F = K \int_{\lambda_1}^{\lambda_2} V_\lambda P_\lambda d\lambda. \quad (A-7)$$

Here P is in watts, and F is in lumens, where 1 lumen = 1 candela-sterad.

Luminous Intensity I:

$$I_\theta = \frac{dF_\theta}{d\omega}. \quad (A-8)$$

The unit of I is the candela.

Luminance L:

$$L_\theta = \frac{dI_\theta}{dA_\theta}. \quad (A-9)$$

The unit of L is the stilb, where 1 stilb = 1 candela/cm².

Illuminance E (Also Called "Illumination"):

$$E = \frac{dF_i}{dA}. \quad (A-10)$$

The unit of E is the phot, where 1 phot = 1 lumen/cm², or the lux, where 1 lux = 1 lumen/m².

Dropping the subscript θ , which indicates the direction in which a quantity is measured, we thus have from Abney's Law for the luminance L of radiation with radiance distribution N_λ ,

$$L = K \int_{\lambda_1}^{\lambda_2} V_\lambda N_\lambda d\lambda, \quad (A-11)$$

where N is measured in watts/sterad-cm², and thus L is in lumens/sterad-cm² = candela/cm² = stilbs.

Retinal Illuminance (The Troland): The troland is a measure of retinal illuminance, being the illuminance produced by viewing a surface of luminance $L = 1$ candela/m² through a pupil opening of 1-mm² area. This unit was formerly called a "photon," but was changed to troland when physicists appropriated photon as the quantum unit.

PHOTOMETRIC UNITS

Candela: The unit of photopic luminous intensity I . The magnitude of the candela is such that the photopic luminance L of a full radiator at the temperature of solidification of platinum is 60 candelas/cm².

Lumen: The unit of photopic luminous flux F . A lumen is the flux emitted in unit solid angle of 1 sterad by a point source having a uniform intensity of 1 candela; hence the total flux from a uniform point source of 1 candela is 4π lumens.

Foot-candle: A unit of illuminance E . A foot-candle is the illuminance of a surface at a distance of 1 ft from a point source of 1-candela intensity; hence, 1 foot-candle = 1 lumen/ft².

Lambert: A unit of luminance L , equal to $(1/\pi)$ candela/cm², i.e., π^{-1} stilb. The lambert is thus equal to the uniform luminance of a perfectly diffusing surface emitting light at the rate of 1 lumen/cm². The foot-lambert (also called the "equivalent foot-candle") is a pseudo-

luminance unit; the luminance L of a perfectly diffusing surface illuminated by E foot-candles is obtained directly in foot-lamberts by multiplying by the luminous reflectance ρ of the surface--

$$L(\text{foot-lamberts}) = \rho E(\text{foot-candles}). \quad (\text{A-12})$$

Conversion Factors:

Illuminance

1 lux (lumen/m ²)	=	0.0929 foot-candle
1 phot (lumen/cm ²)	=	929 foot-candles
1 foot-candle (lumen/ft ²)	=	1 foot-candle

Luminance

1 stilb (candela/cm ²)	=	3.142 lamberts
1 candela/ft ²	=	0.003380 lambert
1 candela/in. ²	=	0.487 lambert
1 foot-lambert	=	0.001076 lambert

BLANK PAGE

REFERENCES

1. Optical Society of America, Committee on Colorimetry, The Science of Color, Optical Society of America, Washington, D.C., 1963.
2. LeGrand, Yves, Light, Color and Vision, Chapman & Hall, Ltd., London, 1957.
3. Evans, R. M., An Introduction to Color, John Wiley & Sons, Inc., New York, 1948.
4. Judd, D. B., Color in Business, Science and Industry, John Wiley & Sons, Inc., New York, 1952.
5. Wright, W. D., The Measurement of Color, The Macmillan Company, New York, 1958.
6. Burnham, R. W., R. M. Hanes, and C. J. Bartleson, Color--A Guide to Basic Facts and Concepts, John Wiley & Sons, Inc., New York, 1963.
7. Dartnall, H. J. A., The Visual Pigments, Methuen & Company, Ltd., London, 1957.
8. Polyak, S. L., The Retina, The University of Chicago Press, Chicago, Ill., 1941.
9. Polyak, S., The Vertebrate Visual System, The University of Chicago Press, Chicago, Ill., 1957.
10. Emsley, H. H., Visual Optics, Hatton Press, Ltd., London, 1953, 2 vols.
11. Zoethout, W. D., Physiological Optics, Professional Press, Inc., Chicago, Ill., 1947.
12. Fulton, J. F., Physiology of the Nervous System, Oxford University Press, New York, 1949.
13. Granit, R., Receptors and Sensory Perception, Yale University Press, New Haven, Conn., 1955.
14. Bartley, S. H., Vision, D. Van Nostrand Company, Inc., New York, 1941.
15. Stevens, S. S. (ed.), Handbook of Experimental Psychology, John Wiley & Sons, Inc., New York, 1951.
16. Koch, S. (ed.), Psychology: A Study of a Science, Vol. I, Sensory, Perceptual, and Physiological Formulations, McGraw-Hill Book Company, Inc., New York, 1959.

17. Osgood, C. E., Method and Theory in Experimental Psychology, Oxford University Press, New York, 1953.
18. Talbot, S. A., "Recent Concepts of Retinal Color Mechanism," J. Opt. Soc. Am., Vol. 41, 1951, p. 895.
19. McMillan, B., "Mathematical Aspects of Information Theory," in Current Trends in Information Theory, University of Pittsburgh Press, Pittsburgh, Pa., 1954, pp. 1-17.
20. Wright, W. D., The Measurement of Color, 3d ed., Hilger & Watts, London, 1964.
21. Optical Society of America, Committee on Colorimetry, The Science of Color, Thomas Y. Crowell Company, New York, 1953.
22. Optical Society of America, Committee on Colorimetry, "The Concept of Color," J. Opt. Soc. Am., Vol. 33, 1943, p. 544.
23. Massachusetts Institute of Technology, Color Measurement Laboratory, Handbook of Colorimetry, A. C. Hardy (ed.), Technology Press, Cambridge, Mass., 1936.
24. Commission Internationale de l'Eclairage, "Minutes of the 6th Session," Geneva, 1924, p. 67, and "Minutes of the Second Plenary Session of the 12th Session," Stockholm, 1951, p. 4.
25. Barbrow, L. E., "Report of the Zurich Meeting of the International Commission on Illumination," J. Opt. Soc. Am., Vol. 45, 1955, p. 894.
26. Judd, D. B., "Basic Correlates of the Visual Stimulus," in Handbook of Experimental Psychology, S. S. Stevens (ed.), John Wiley & Sons, Inc., New York, 1951.
27. Wright, W. D., "A Re-determination of the Trichromatic Coefficients of the Spectral Colors," Trans. Opt. Soc. London, Vol. 30, 1928-1929, p. 141; also Vol. 31, 1929-1930, p. 201.
28. Guild, J., "The Colorimetric Properties of the Spectrum," Phil. Trans. Roy. Soc. London, Ser. A, Vol. 230, 1931, p. 149.
29. Kelly, K. L., "Color Designations for Lights," J. Res. Nat. Bur. Std., Vol. 31, 1943, p. 271.
30. Judd, D. B., "Hue, Saturation and Lightness of Surface Colors with Chromatic Illumination," J. Opt. Soc. Am., Vol. 30, 1940, p. 2.
31. Evans, R. M., "Variables of Perceived Color," J. Opt. Soc. Am., Vol. 54, 1964, p. 1467.

32. Judd, D. B., "Radical Changes in Photometry and Colorimetry Fore-shadowed by C.I.E. Actions in Zurich," J. Opt. Soc. Am., Vol. 45, 1955, p. 897.
33. Barbrow, L. E., "Brussels Session of the International Commission on Illumination," J. Opt. Soc. Am., Vol. 50, 1960, p. 89.
34. Wald, G., "The Photochemistry of Vision," Doc. Ophthalmol., Vol. 3, 1949, p. 94.
35. Walters, H. V., and W. D. Wright, "The Spectral Sensitivity of the Fovea and Extrafovea in the Purkinje Range," Proc. Roy. Soc. (London), Ser. B, Vol. 131, 1943, p. 340.
36. Wald, G., "Human Vision and the Spectrum," Science, Vol. 101, 1945, p. 653.
37. Gibson, K. S., and E. P. Tyndall, "Visibility of Radiant Energy," Sci. Papers Nat. Bur. Std., Vol. 19, 1925, p. 131.
38. Gibson, K. S., "Spectral Luminosity Factors," J. Opt. Soc. Am., Vol. 30, 1940, p. 51.
39. Kinney, J. A. S., "Sensitivity of the Eye to Spectral Radiation at Scotopic and Mesopic Intensity Levels," J. Opt. Soc. Am., Vol. 45, 1955, p. 507.
40. Kinney, J. A. S., "Comparison of Scotopic, Mesopic, and Photopic Spectral Sensitivity Curves," J. Opt. Soc. Am., Vol. 48, 1958, p. 185.
41. Weaver, K. S., "The Visibility of Radiation at Low Intensities," J. Opt. Soc. Am., Vol. 27, 1937, p. 36.
42. Hecht, S., and R. E. Williams, "The Visibility of Monochromatic Radiation," J. Gen. Physiol., Vol. 5, 1922, p. 1.
43. Sloan, L. L., "The Effect of Intensity, Adaptation State, and Field Size on the Visibility Curve," Psych. Monographs, Vol. 38, 1928, No. 1.
44. Grigorovici, R., and I. Aricescu-Savopol, "Luminosity and Chromaticity in the Mesopic Range," J. Opt. Soc. Am., Vol. 48, 1958, p. 891.
45. Hsia, Y., and C. H. Graham, "Spectral Sensitivity of the Cones in the Dark Adapted Human Eye," Proc. Nat. Acad. Sci. U.S., Vol. 38, 1952, p. 80.
46. Sperling, H. G., and Y. Hsia, "Some Comparisons among Spectral Sensitivity Data Obtained in Different Retinal Locations and with Two Sizes of Foveal Stimulus," J. Opt. Soc. Am., Vol. 47, 1957, p. 707.

47. Boynton, R. M., G. Kandel, and J. W. Onley, "Rapid Chromatic Adaptation of Normal and Dichromatic Observers," J. Opt. Soc. Am., Vol. 49, 1959, p. 654.
48. Thomson, L. C., "Shape Irregularities in the Equal Energy Luminosity Curve," Proc. Phys. Soc. (London), Vol. B62, 1949, p. 787.
49. Bedford, R. E., and G. W. Wyszecki, "Luminosity Functions for Various Field Sizes and Levels of Retinal Illuminance," J. Opt. Soc. Am., Vol. 48, 1958, p. 406.
50. Weale, R. A., "The Foveal and Para-central Spectral Sensitivities in Man," J. Physiol. (London), Vol. 114, 1951, p. 435.
51. Ishak, I. G. H., "The Photopic Luminosity Curve for a Group of Fifteen Egyptian Trichromats," J. Opt. Soc. Am., Vol. 42, 1952, p. 529.
52. Hurvich, L. M., and D. Jameson, "Spectral Sensitivity of the Fovea, I, Neutral Adaptation," J. Opt. Soc. Am., Vol. 43, 1953, p. 485.
53. Stiles, W. S., "Interim Report to the C.I.E., Zurich, 1955, on the N.P.L. Investigation of Colour Matching," Opt. Acta, Vol. 2, 1955, p. 168.
54. Willmer, E. N., "Further Observations on the Properties of the Central Fovea in Colour Blind and Normal Subjects," J. Physiol. (London), Vol. 110, 1949, p. 422.
55. Thomson, L. C., "The Spectral Sensitivity of the Central Fovea," J. Physiol. (London), Vol. 112, 1951, p. 114.
56. LaDriere, M., "Variations in Retinal Quantum Absorptions of the Fovea as a Function of Wavelength and Area," J. Psych., Vol. 51, 1961, p. 213.
57. Collins, W. E., "The Effects of Deuteranomaly and Deuteranopia upon the Foveal Luminosity Curve," J. Psych., Vol. 48, 1959, p. 285.
58. Collins, W. E., "Luminosity Functions of Normal, Deuteranomalous, and Deuteranopic Subjects as Determined by Absolute Threshold and CFF Measurements," J. Opt. Soc. Am., Vol. 51, 1961, p. 202.
59. Hsia, Y., and C. H. Graham, "Spectral Luminosity Curves for Protanopic, Deuteranopic, and Normal Subjects," Proc. Nat. Acad. Sci. U.S., Vol. 43, 1957, p. 1011.
60. Sperling, H. G., and W. G. Lewis, "Some Comparisons between Foveal Spectral Sensitivity Data Obtained at High Brightness and Absolute Threshold," J. Opt. Soc. Am., Vol. 49, 1959, p. 983.

61. Heath, G. G., "Luminosity Curves of Normal and Dichromatic Observers," Science, Vol. 128, 1958, p. 775.
62. Sperling, H. G., and C. L. Joliffe, "Chromatic Response Mechanisms in the Human Fovea as Measured by Threshold Spectral Sensitivity," Science, Vol. 136, 1962, p. 317.
63. Boynton, R. M., "Rapid Chromatic Adaptation and the Sensitivity Functions of Human Color Vision," J. Opt. Soc. Am., Vol. 46, 1956, p. 172.
64. Jameson, D., and L. M. Hurvich, "Spectral Sensitivity of the Fovea, II, Dependence on Chromatic Adaptation," J. Opt. Soc. Am., Vol. 43, 1953, p. 552.
65. Wright, W. D., "Spectral Sensitivity of the Retinal Receptors," Nature, Vol. 151, 1943, p. 726.
66. Dillon, D. J., and R. T. Zegers, "Quantal Determination and Statistical Evaluation of Absolute Foveal Luminosity Thresholds and of Threshold Variability," J. Opt. Soc. Am., Vol. 48, 1958, pp. 877-883.
67. Jackson, Janice M., "Variability in Visual Thresholds," J. Psych., Vol. 59, 1965, pp. 17-28.
68. Zegers, Richard T., private communication, January 1965.
69. Hurvich, L. M., and D. Jameson, "Spectral Sensitivity of the Fovea, III, Heterochromatic Brightness and Chromatic Adaptation," J. Opt. Soc. Am., Vol. 44, 1954, p. 213.
70. Willmer, E. N., and W. D. Wright, "Colour Sensitivity of the Fovea Centralis," Nature, Vol. 156, 1945, p. 119.
71. Hartridge, H., "The Visual Perception of Fine Detail," Phil. Trans. Roy. Soc. London, Ser. B, Vol. 232, 1947, p. 592.
72. Hartridge, H., "Recent Advances in Colour Vision," Advan. Sci., Vol. 5, 1949-1959, p. 243.
73. Stiles, W. S., "A Modified Helmholtz Line Element in Brightness-Color Space," Proc. Phys. Soc. (London), Vol. 58, 1946, p. 41.
74. Stiles, W. S., and J. M. Burch, "N.P.L. Color-matching Investigation: Final Report (1958)," Opt. Acta, Vol. 6, 1959, p. 1.
75. Stiles, W. S., "Current Problems of Visual Research," Proc. Phys. Soc. (London), Vol. 56, 1944, p. 329; excerpted in Nature, Vol. 154, 1944, p. 290.

76. Wald, George, "The Receptors of Human Color Vision," Science, Vol. 145, 1964, p. 1007.
77. Boynton, R. M., "Contributions of Threshold Measurements to Color-Discrimination Theory," J. Opt. Soc. Am., Vol. 53, 1963, pp. 165-178.
78. Stiles, W. S., "Color Vision: The Approach through Increment-Threshold Sensitivity," Proc. Nat. Acad. Sci. U.S., Vol. 45, 1959, pp. 100-114.
79. Ikeda, Mitsuo, "Study of Interrelations between Mechanisms at Threshold," J. Opt. Soc. Am., Vol. 53, 1963, pp. 1305-1313.
80. Boynton, R. M., Mitsuo Ikeda, and W. S. Stiles, "Interactions among Chromatic Mechanisms as Inferred from Positive and Negative Increment Thresholds," Vision Res., Vol. 4, 1964, pp. 87-117.
81. Moon, P., and D. E. Spencer, "Analytical Representation of Standard Response Curves," J. Opt. Soc. Am., Vol. 33, 1943, p. 89.
82. Crescitelli, F., and H. J. A. Dartnall, "Human Visual Purple," Nature, Vol. 172, 1953, p. 195.
83. Wald, G., and P. K. Brown, "Human Rhodopsin," Science, Vol. 127, 1958, p. 222.
84. Crescitelli, F., "The Natural History of Visual Pigments," Ann. New York Acad. Sci., Vol. 74, 1958, p. 230.
85. Wald, G., "The Chemistry of Rod Vision," Science, Vol. 113, 1951, p. 287.
86. Moon, P., and D. E. Spencer, "A Modified Photochemical Theory of Vision," J. Opt. Soc. Am., Vol. 35, 1945, p. 430.
87. Kropf, A., and R. Hubbard, "The Mechanism of Bleaching Rhodopsin," Ann. New York Acad. Sci., Vol. 74, 1958, p. 266.
88. Weale, R. A., "Further Studies of Photo-Chemical Reactions in Living Human Eyes," Vision Res., Vol. 1, 1962, p. 354.
89. Dartnall, H. J. A., and C. F. Goodeve, "Scotopic Luminosity Curve and the Absorption Spectrum of Visual Purple," Nature, Vol. 139, 1937, p. 409.
90. Ludvigh, E., and E. F. McCarthy, "Absorption of Visible Light by the Refractive Media of the Human Eye," Arch. Ophthalmol., Vol. 20, 1938, p. 37.

91. Rushton, W. A. H., and F. W. Campbell, "Measurement of Rhodopsin in the Living Human Eye," Nature, Vol. 174, 1954, p. 1096.
92. Brown, P. K., and G. Wald, "Visual Pigments in Human and Monkey Retinas," Nature, Vol. 200, 1963, p. 37.
93. Brown, P. K., and G. Wald, "Visual Pigments in Single Rods and Cones of the Human Retina," Science, Vol. 144, 1964, p. 45.
94. Wald, G., "Photo-labile Pigments of the Chicken Retina," Nature, Vol. 140, 1937, p. 545.
95. Dartnall, H. J. A., "Visual Pigment from a Pure-Cone Retina," Nature, Vol. 188, 1960, p. 475.
96. Weale, R. A., "Bleaching Experiments on Eyes of Living Gray Squirrels," J. Physiol. (London), Vol. 127, 1955, p. 587.
97. Rushton, W. A. H., "Physical Measurement of Cone Pigment in the Living Human Eye," Nature, Vol. 179, 1957, p. 571.
98. Rushton, W. A. H., "Kinetics of Cone Pigments Measured Objectively on the Living Human Fovea," Ann. New York Acad. Sci., Vol. 74, 1958, p. 291.
99. Ripps, H., and R. A. Weale, "On Seeing Red," J. Opt. Soc. Am., Vol. 54, 1964, p. 272.
100. Rushton, W. A. H., "Interpretation of Retinal Densitometry," J. Opt. Soc. Am., Vol. 54, 1964, p. 273.
101. Rushton, W. A. H., "A Cone Pigment in the Protanope," J. Physiol. (London), Vol. 168, 1963, p. 345.
102. Rushton, W. A. H., "The Density of Chlorolabe in the Foveal Cones of the Protanope," J. Physiol. (London), Vol. 168, 1963, p. 360.
103. Rushton, W. A. H., "Cone Pigment Kinetics in the Protanope," J. Physiol. (London), Vol. 168, 1963, p. 374.
104. Marks, W. B., W. H. Dobelle, and E. F. MacNichol, Jr., "Visual Pigments of Single Primate Cones," Science, Vol. 143, 1964, p. 1181.
105. Marks, W. B., "Visual Pigments of Single Goldfish Cones," J. Physiol. (London), Vol. 178, 1965, p. 33.
106. Liebman, P. A., and G. Entine, "Sensitive Low-light-level Microspectrophotometer: Detection of Photosensitive Pigments of Retinal Cones," J. Opt. Soc. Am., Vol. 54, 1964, pp. 1451-1459.

107. Snitzer, E., and H. Osterberg, "Observed Dielectric Waveguide Modes in the Visible Spectrum," J. Opt. Soc. Am., Vol. 51, 1961, p. 499.
108. Enoch, J. M., "Nature of the Transmission of Energy in the Retinal Receptors," J. Opt. Soc. Am., Vol. 51, 1961, p. 1122.
109. Marks, William B., private communication, January 1965.
110. De Valois, R. L., "Color Vision Mechanisms in the Monkey," J. Gen. Physiol., Vol. 43, 1960, pp. 115-128.
111. Boynton, R. M., "Discussion: Competing Theories of Receptor Excitation," Psych. Bull., Vol. 61, 1964, pp. 262-267.
112. Vilter, V., "Recherches Biometriques sur l'Organisation Synaptique de la Retine Humaine," Comp. Rend. Soc. Biol., Vol. 143, 1949, p. 830.
113. Barer, R., "Refractometry and Interferometry of Living Cells," J. Opt. Soc. Am., Vol. 47, 1957, p. 545.
114. O'Brien, B., "Vision and Resolution in the Central Retina," J. Opt. Soc. Am., Vol. 41, 1951, p. 882.
115. Dowling, J. E., "Foveal Receptors of the Monkey Retina: Fine Structure," Science, Vol. 147, 1965, pp. 57-59.
116. Dowling, John E., private communication, January 1965.
117. Enoch, J. M., "Waveguide Modes: Are They Present, and What Is Their Possible Role in the Visual Mechanism?" J. Opt. Soc. Am., Vol. 50, 1960, p. 1025.
118. Kelsey, P. A., and I. Schwartz, "Nature of the Limit of the Color Zone in Perimetry," J. Opt. Soc. Am., Vol. 49, 1959, p. 764.
119. Snitzer, B., "Cylindrical Dielectric Waveguide Modes," J. Opt. Soc. Am., Vol. 51, 1961, p. 491.
120. Noell, W. K., "Studies on Visual Cell Viability and Differentiation," Ann. New York Acad. Sci., Vol. 74, 1958, p. 337.
121. Hecht, S., and S. Shlaer, "The Relation between Intensity and Critical Frequency for Different Parts of the Spectrum," J. Gen. Physiol., Vol. 19, 1936, p. 965.
122. Brindley, G. S., "Beats Produced by Simultaneous Stimulation of the Human Eye with Intermittent Light and Intermittent or Alternating Electric Current," J. Physiol. (London), Vol. 164, 1962, p. 157.

123. Granit, R., "A Physiological Theory of Colour Perception," Nature, Vol. 151, 1943, p. 11.
124. Granit, R., Sensory Mechanisms of the Retina, Oxford University Press, New York, 1947.
125. Donner, K. O., "The Spectral Sensitivity of the Pigeon's Retinal Elements," J. Physiol. (London), Vol. 122, 1953, p. 524.
126. Jahn, T. L., "Color Vision and Color Blindness: A Mechanism in Terms of Modern Evidence," J. Opt. Soc. Am., Vol. 36, 1946, p. 595.
127. Chapman, R. M., "Spectral Sensitivity of Single Neural Units in the Bullfrog Retina," J. Opt. Soc. Am., Vol. 51, 1961, p. 1102.
128. Svaetichin, G., and E. F. MacNichol, Jr., "Retinal Mechanisms for Chromatic and Achromatic Vision," Ann. New York Acad. Sci., Vol. 74, 1958, p. 385.
129. Rosenberg, B., R. J. Heck, and K. Aziz, "Color Responses in an Organic Photoconductive Cell," J. Opt. Soc. Am., Vol. 54, 1964, p. 1018.
130. MacNichol, E. F., Jr., "Retinal Mechanisms of Color Vision," Vision Res., Vol. 4, 1964, pp. 119-133.
131. Smirnov, G. D., "Comparative Approach to the Neurophysiology of Vision," in Brain and Behavior, Vol. 1, M. Brazier (ed.), American Institute of Biological Sciences, Washington, D.C., 1961.
132. Clark, W. E. LeGros, "The Laminar Pattern of the Lateral Geniculate Nucleus Considered in Relation to Colour Vision," Doc. Ophthalmol., Vol. 3, 1949, p. 57.
133. Hect, S., "On the Binocular Fusion of Colors and Its Relation to Theories of Color Vision," Proc. Nat. Acad. Sci. U.S., Vol. 14, 1928, p. 237.
134. Pickford, R. W., "Binocular Colour Combinations," Nature, Vol. 159, 1947, p. 268.
135. DeValois, R. L., C. J. Smith, S. T. Kitai, and A. J. Karoly, "Response of Single Cells in Monkey Lateral Geniculate Nucleus to Monochromatic Light," Science, Vol. 127, 1958, pp. 238-239.
136. DeValois, R. L., C. J. Smith, A. J. Karoly, and S. T. Kitai, "Electrical Responses of Primate Visual System: I. Different Layers of Macaque Lateral Geniculate Nucleus," J. Comp. Physiol. Psych., Vol. 51, 1958, pp. 662-668.

137. Smith, C. J., R. L. DeValois, and S. T. Kitai, "Microelectrode Recordings from Cells in Visual Cortex of Macaques," Am. Psych., Vol. 13, 1958, p. 387.
138. Hartridge, H., "Recent Advances in Color Vision," Science, Vol. 108, 1948, p. 395.
139. Holmes, G., "The Organization of the Visual Cortex in Man," Proc. Roy. Soc. (London), Ser. B, Vol. 132, 1945, p. 348.
140. Haider, M., P. Spong, and D. B. Lindsley, "Attention, Vigilance, and Cortical Evoked-Potentials in Humans," Science, Vol. 145, 1964, p. 180.
141. Spong, Paul, M. Haider, and D. B. Lindsley, "Selective Attentiveness and Cortical Evoked Responses to Visual and Auditory Stimuli," Science, Vol. 148, 1965, pp. 395-397.
142. Davis, H., "Enhancement of Evoked Cortical Potentials in Humans Related to a Task Requiring a Decision," Science, Vol. 145, 1964, p. 182.
143. Burnham, R. W., R. M. Evans, and S. M. Newhall, "Influence on Color Perception of Adaptation to Illumination," J. Opt. Soc. Am., Vol. 42, 1952, p. 597.
144. Burnham, R. W., "Predictions of Shifts in Color Appearance with a Change from Daylight to Tungsten Adaptation," J. Opt. Soc. Am., Vol. 49, 1959, p. 254.
145. MacAdam, D. L., "Chromatic Adaptation," J. Opt. Soc. Am., Vol. 46, 1956, p. 500.
146. Hurvich, L. M., and D. Jameson, "A Psychophysical Study of White, I, Neutral Adaptation," J. Opt. Soc. Am., Vol. 41, 1951, p. 521.
147. Jameson, D., and L. M. Hurvich, "A Psychophysical Study of White, II, Neutral Adaptation, Area and Duration as Variants," J. Opt. Soc. Am., Vol. 41, 1951, p. 528.
148. Jameson, D., and L. M. Hurvich, "A Psychophysical Study of White, III, Adaptation as Variant," J. Opt. Soc. Am., Vol. 41, 1951, p. 787.
149. Kohler, J., "Über Aufbau und Wandlungen der Wahrnehmungswelt," Akad. Wiss., Vienna: Philosoph. - Hist. Klasse - Sitzber., Vol. 227, 1951, pp. 1-118. (Translated by Harry Fiss as "The Formation and Transformation of the Perceptual World," Psych. Issues, Vol. III, No. 4, Monograph 12, 1964.)

150. Land, E. H., "Color Vision and the Natural Image," Proc. Nat. Acad. Sci. U.S., Vol. 45, 1959, pp. 115, 636.
151. Walls, G. L., "Land! Land!" Psych. Bull., Vol. 57, 1960, p. 29.
152. Judd, D. B., "Appraisal of Land's Work on Two-Primary Color Projections," J. Opt. Soc. Am., Vol. 50, 1960, p. 254.
153. Woolfson, M. M., "Some New Aspects of Color Perception," IBM J., October 1959, p. 312.
154. Siegel, M. H., and F. L. Dimmick, "Discrimination of Color, II, Sensitivity as a Function of Spectral Wavelength, 510 to 630 mμ," J. Opt. Soc. Am., Vol. 52, 1962, p. 1071.
155. Campbell, J. W., "Color-Vision--The Land Experiments," Astounding Sci. Fact and Fiction, Vol. 64, February 1960, p. 83.
156. Land, E. H., "The Retinex," Am. Scientist, Vol. 52, 1964, p. 247.
157. Helson, H., "Some Factors and Implications of Color Constancy," J. Opt. Soc. Am., Vol. 33, 1953, p. 555.
158. Wright, W. D., "The Maxwell Colour Centenary," Nature, Vol. 191, 1961, p. 10.
159. Helson, H., "Adaptation Level Theory," in Psychology: A Study of a Science, Vol. I, Sensory, Perceptual, and Physiological Formulations, S. Koch (ed.), McGraw-Hill Book Company, Inc., New York, 1959, pp. 565-621.

DOCUMENT CONTROL DATA

1. ORIGINATING ACTIVITY THE RAND CORPORATION		2a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED
		2b. GROUP
3. REPORT TITLE A CRITICAL REVIEW OF THE EXPERIMENTAL FOUNDATION OF HUMAN COLOR PERCEPTION		
4. AUTHOR(S) (Last name, first name, initial) Sheppard, Jr., Joseph J.		
5. REPORT DATE January 1966	6a. TOTAL NO. OF PAGES 187	6b. NO. OF REFS. 159
7. CONTRACT or GRANT NO. SD-79	8. ORIGINATOR'S REPORT NO. RM-4196-ARPA	
9a. AVAILABILITY/LIMITATION NOTICES DDC 1		9b. SPONSORING AGENCY Advanced Research Projects Agency
10. ABSTRACT <p>A presentation of the minimum material needed for a comprehensive study of normal human color perception. The artificial nature of colorimetry is discussed, with emphasis on the distinction between experimental facts established in the matching experiments and the formalism of colorimetry derived in part from these facts. A representative portion of the available experimental data on individual foveal spectral sensitivity is collected and analyzed. Data on the initial photoreception process and the anatomy, histology, morphology, ontogeny, and electrophysiology of the retinal neurons are analyzed, noting the multiplicity of results indicating a fundamental difference between receptor mechanisms in the rods and cones. A review of data on central neural mechanisms indicates a complex, dynamic role for the lateral geniculate nuclei in human color vision. Considerations of diverse psychophysiological phenomena are summarized. The general conclusion of the study is that the available experimental evidence does not clearly dictate the fundamental physiological processes mediating human color vision. Principal conclusions are discussed in relation to the three distinct fields of colorimetry, visual biophysics, and visual psychophysics. Four suggestions are given for psychophysical modelling.</p>		11. KEY WORDS Color Physics Physiology Vision Psychology Neurophysiology